Induced Charge Computation

Dissertation
zur Erlangung der Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Goethe-Universität Frankfurt am Main

von
Mehmet Süzen
aus
Zypern

Frankfurt am Main
2009
(D-30)

vom Fachbereich Physik der Goethe-Universität
Frankfurt am Main als Dissertation angenommen.

Dekan : Prof. Dr. Dirk-Hermann Rischke

Gutachter 1: PD. Dr. Christian Holm
Gutachter 2: JP. Dr. Karin Hauser

Datum der Disputation: 03.07.2009

In loving memory of
my father *Teoman Süzen* (1943 - 2006)
my grandfather UK SBA Insp. *M. Kalfaoğlu* (1925 - 2008)
my aunth *Sema Erüreten* (1953 - 2007)
my cousin national athlete *M. Ergazi* (1986 - 2006)

Summary

One of the main aspects of statistical mechanics is that the properties of a thermodynamics state point do not depend on the choice of the statistical ensemble. It breaks down for small systems e.g. single molecules. Hence, the choice of the statistical ensemble is crucial for the interpretation of single molecule experiments, where the outcome of measurements depends on which variables or control parameters, are held fixed and which ones are allowed to fluctuate. Following this principle, this thesis investigates the thermodynamics of a single polymer pulling experiments within two different statistical ensembles.

The scaling of the conjugate chain ensembles, the fixed end-to-end vector (Helmholtz) and the fixed applied force (Gibbs), are studied in depth. This thesis further investigates the ensemble equivalence for different force regimes and polymer-chain contour lengths. Using coarse-grained molecular dynamic simulations, i.e. Langevin dynamics, the simulations were found to complement the theoretical predictions for the scaling of ensemble difference of Gaussian chains in different force-regimes, giving special attention to the zero force regime.

After constructing Helmholtz and Gibbs conjugate ensembles for a Gaussian chain, two different data sets of thermodynamic states on the force-extension plane, i.e. force-extension curves, were generated. The ensemble difference is computed for different polymer-chain lengths by using force-extension curves. The scaling of the ensemble difference versus relative polymer-chain length under different force regimes has been derived from the simulation data and compared to theoretical predictions. The results demonstrate that the Gaussian chain in the zero force limit generates nonequivalent ensembles, regardless of its equilibrium bond length and polymer-chain contour length.

Moreover, if polymers are charged in confinement, coarse-graining is problematic, owing to dielectric interfaces. Hence, the effect of dielectric interfaces must be taken into account when describing physical systems such as ionic channels or biopolymers inside nanopores. It is shown that the effect of dielectrics is crucial for the dynamics of a biopolymer or an ion inside a nanopore. In the simulations, the feasibility of an efficient and accurate computation of electrostatic interactions in the presence of an arbitrarily shaped dielectric domain is challenging. Several solutions for this problem have been previously proposed in the literature such as a density functional approach, or transforming problem at hand into an algebraic problem (*Induced Charge Computation* (ICC)) and boundary element methods. Even though the essential concept is the same, which is to replace the dielectric interface with a polarization charge density, these approaches have been analyzed and the ICC algorithm has been implemented. A new superior boundary element method has been devised utilizing the force computation via the Particle-Particle Particle-Mesh (P^3M) method for periodic geometries ($ICCP^3M$). This method has been compared to the ICC algorithm, the algebraic solutions, and to density functional approaches. Extensive numerical tests against analytically tractable geometries have confirmed the correctness and applicability of developed and implemented algorithms, demonstrating that the $ICCP^3M$ is the fastest and the most versatile algorithm. Further optimization issues are also discussed in obtaining accurate induced charge densities.

The potential of mean force (PMF) of DNA modelled on a coarsed-grain level inside a nanopore is investigated with and without the inclusion of dielectric effects. Despite the simplicity of the model, the dramatic effect of dielectric inclusions is clearly seen in the observed force profile.

Kurzfassung

Eines der wichtigsten Ergebnisse der statistischen Mechanik ist, dass unterschiedliche statistische Ensembles dieselben thermodynamischen Zustände erzeugen. Dieses Prinzip gilt nicht notwendigerweise für kleine Systeme, wie zum Beispiel einzelne Moleküle oder ein einzelnes Polymer. Deshalb ist die Wahl des statistischen Ensembles von entscheidender Bedeutung für die Interpretation von Einzelmolekülexperimenten (im Englischen "Single Molecule Experiment" (SME)), denn das Ergebnis der Messung hängt davon ab, welche Variablen oder Kontrollparameter festgehalten werden und welche fluktuieren können.

Ausgehend von diesem Problem haben wir Zugexperimente an einem einzigen Polymer in zwei verschiedenen Ensembles durchgeführt und den thermodynamischen Limes (Anzahl der Polymersegmente wächst gegen unendlich) untersucht. Wir haben zwei konjugierte Ensembles, nämlich das, in dem der End-zu-End Abstand (Helmholtz) festgehalten wurde, mit dem, wo wir die Kraft (Gibbs) festgehalten haben, gründlich und auf verschiedene Arten verglichen. Wir haben den Ensemble-Unterschied als Funktion der Anzahl der Polymersegmente in unterschiedlichen Zugkraftbereichen mittels Molekulardynamik Simulationen untersucht, wobei wir eine Langevin Dynamik benutzt haben. Die untersuchten Messgrössen waren die Bestimmung von sogenannten Kraft-Dehnungskurven, wie sie auch in AFM Experimenten gemessen werden. Diese Kurven wurden für zwei verschieden Gauss Ketten verschiedenster Polymerlänge durchgeführt, einmal mit verschwindender Bondlänge und einmal mit Bondlänge eins.

Aufgrund unserer Simulationen konnten wir zeigen, das sowohl Gauss-Ketten mit endlicher, wie auch verschwindender Bondlänge für den Bereich verschwindender Zugkraft einen endlichen Ensembleunterschied besitzen, der nicht von der Kettenlänge abhängt. Dieses Phänomen wurde bereits vor 20 Jahren von R. Neumann beschrieben. Trotz der relativ einfachen Argumente von Neumann gibt es bis heute noch Arbeiten, die diesen Sachverhalt entweder anzweifeln oder verkehrt darstellen. Wir hoffen, durch diesen Teil der Arbeit den Sachverhalt zufriedenstellend aufgeklärt zu haben.

Im zweiten Teil der Arbeit behandeln wir geladen Polymere unter einem räumlichen Einschluss. Dies können zum Beispiel Ionen in schmalen Kanälen sein (Ionenkanäle), oder DNA in Nanoporen. In vergröberten Simulationen werden geladene Polymere immer in einem dielektrischen Kontinuum dargestellt. Wasser hat eine relative dielektrische Konstante von 80 bei Raumtemperatur, die dann in dieses Model als Parameter gesteckt wird. Wenn feste Grenzflächen vorhanden sind, haben diese meist niedrige dielektrische Konstanten (≈ 2). Diese Grenzflächen haben grosse Auswirkungen auf die elektrostatischen Wechselwirkungen. In den Simulationen ist es wichtig, diese Effekte korrekt **und schnell** zu berechnen. Deshalb haben wir einen effizienten und präzisen Algorithmus entwickelt, der genau dies bewerkstelligt. In der Literatur wurden mehrere Möglichkeiten vorgeschlagen, wie dieses Problem für Simulationen lösbar sein sollte, wie zum Beispiel Dichtefunktionalmethoden, Umwandlung des Problems in ein algebraisches Problem (Induced Charge Computation, ICC) oder die Randelement Methode. Das wesentliche Konzept besteht darin, die Polarisationsladung auf dem dielektrischen Rand so zu bestimmen, dass die dielektrischen Randbedingungen erfüllt werden. Wir haben den ICCP^{3}M Algorithmus entwickelt,

dessen Kernstück darin besteht, den P^{3}M Algorithmus zur Bestimmung der induzierten Ladung auf den Randelementen zu benutzen. Durch diesen Trick lässt sich die Ladungsberechnung in CPU Zeit $\mathcal{O}(NlogN)$, wobei $\mathcal{O}(N)$ die Anzahl der Ladungen im System ist, durchführen. Wir haben den Algorithmus innerhalb des Espresso Programmpakets implementiert und optimiert.

Im letzten Teil der Arbeit wurde das Potential der mittleren Kraft einer vergröberten DNA innerhalb einer Nanopore untersucht, wobei wir die Unterschiede zwischen korrekter Behandung der dielektrischen Ränder und der Ignorierung derselben quantifiziert haben. Trotz seiner Einfachheit zeigt unser Modell den dramatischen Einfluss, den die dielektrischen Ränder auf die gemessene efffektive Kraft und das Potential der mittleren Kraft ausüben.

Contents

List of Figures

1 Introduction

Biology and Physics are normally regarded as two distinct areas of knowledge where the former deals with living organisms and their interaction with the environment as well as with each other, while the later is the science that deals with the structure of matter and the behavior of the universe (for all practical purposes considered as non-living). Nevertheless, in the last decade or so both the fields have converged and generated overlapping research areas, at the experimental as well as theoretical levels, in a revolutionary way. For example, the ability to manipulate single molecules experimentally [1, 2, 3] has become one of the central paradigms in today's soft condensed matter physics and biophysics research. These manipulation techniques, collectively known as Single Molecule Experiment (SME) [2], in particular combined with the presence of natural or synthetic nanopores [4, 5, 6], are necessary to explain transport phenomena within the framework of nanofluidics [7] as well as in natural biological processes [8, 9, 1, 2]. Some of the notable techniques are the atomic force microscopy and optical or magnetic tweezers, which enable us to study the mechanical properties of single macromolecules, such as DNA, exposed *in vivo* to stretching forces [10]. Information from these SMEs is usually obtained in the form of force-extension curves (FEC) [10]. The FEC quantitatively explains the behaviour of the macromolecular extension (end-to-end distance) upon force applied to one end of a macromolecule, while the other end is held fixed.

The SMEs are a particularly important experimental tool to understand the statistical mechanics of small systems, where thermal fluctuations are dominant. Hence, careful redefinition of thermodynamics for these systems is necessary [11, 12, 13, 14, 15, 16]. In the same context, the SMEs constitute a fundamental problem: within the limit of a single short polymer, the outcome of measurements may depend upon which observables are held fixed and which are allowed to fluctuate [17]. This phenomenon is related to the problem of statistical ensemble inequivalance in polymer stretching experiments [18, 19, 20, 21, 22, 23, 24, 25, 26]. The experimental setting on short polymer can be either, *isotensional*, where the pulling force is fixed and the extension is fluctuating, or *isometric*, where the extension is fixed and the force is fluctuating [17]. This type of behavior has been investigated in the context of the statistical mechanics of elasticity [27], the theoretical analysis has been given in references [28, 29, 17, 30]. Yet, to the best of our knowledge there are no simulation studies available, that compare the outcomes of two polymer ensembles to check their equivalence in depth, particularly within a wide range of force regimes. This thesis provides molecular simulation results on the equivalence of two polymer stretching ensembles. Specifically, the question that is addressed is whether ensemble equivalence in the thermodynamic limit is possible in the low stretching force regimes by quantitatively determining the scaling of ensemble difference against the polymer length.

One novel SME force measurements combined with the presence of synthetic nanopore is based on the idea of pulling a DNA molecule out of a nanopore with Optical Tweezers [31, 32]. This kind of approach open a new dimension in the understanding of DNA electrostatic properties and allows scientists to measure the effective charge of DNA directly. In similar contexts, these types of experiments could give rise to new perspectives in DNA sequencing [33, 34, 35]. However, we

would like to model a similar situation with coarse-grained simulations.

Computer simulations of soft materials and biophysical systems as a sub-field of computational physics and chemistry have also attracted recent attention [36]. Due to spatial and time scale restrictions in full atomistic simulations of soft matter and biophysical processes, coarse-grained models have become the standard investigation tools [36, 37, 38]. Similarly, in order to mimic the above mentioned SME experiment with coarse-grained molecular simulations, a model for DNA and its environment has to be constructed, such as the one in similar DNA condensation studies [39, 40]. In the present model, overlapping excluded volume interactions have been introduced, along with DNA backbone charges and its counter ions. In the environment, the water and synthetic pore (a channel) embedded in a polarizable membrane are going to be modeled with continuum electrostatics, having dielectric constants of 80 and 2 respectively.

A different problem arises, if charged polymers are located in the confinement, such as a polarizable membrane like the one discussed above; the effect of dielectric interfaces has to be taken into account. The striking effect of the electric field in this situation appears due to the presence of the dielectric jumps between the confining medium and the aqueous solution [41]. This effect usually comes with the force generated by charges induced in the dielectric interface which is called Dielectric Boundary Force (DBF) [42, 43, 44]. In coarse-grained dynamics of a macromolecule that translocates through pores, computing the DBF for each time step is a challenging task given its computational complexity. To determine the DBF one also needs to solve Poisson equation for inhomogeneous dielectrics, which transforms into Poisson-Boltzmann (PB) equation if an ionic atmosphere is added to the system [45], or into Poisson-Nernst-Planck (PNP) equation if charge current is taken into account. Solving these non-linear partial differential equations is therefore not straightforward.

Although the problem of treating different dielectrics has been investigated in depth in biophysics [46, 47, 48, 49], chemistry [45, 50, 51] computational mathematics [52, 53], electronic device simulations [54], and physics [55, 56], the careful analysis of these references has shown that every known solution lacks required features, which are computational efficiency, accuracy and the range of applicability. A novel way of computing the DBF is to represent the effect of an interfacial dielectric as a polarization charge density on the dielectric boundary [48, 57]. This proposed method to solve the above mentioned boundary value problem is known as Induced Charge Computation (ICC) [58, 59], where mathematically the problem boils down to solving the set of linear equations for induced charge density. The ICC has been implemented, and this method has introduced severe limitations in the number of induced charges that can be treated, due to excessive computational memory requirements; it is also lacking the proper treatment of periodic boundary conditions.

We have proposed an alternative solution to the problem, using the iterative approach to dielectrics and coupling its force component with a Poisson solver in periodic boundaries, called Particle-Particle Particle Mesh, P^3M algorithm [60]. The hybrid method, $ICCP^3M$ has been implemented within ESPResSo [61]. The test results yield the correct polarization charges on the boundary surface for test geometries, where analytical solutions are known. In addition, comparisons have been made of $ICCP^3M$ and ICC on the test system.

Furthermore, the new algorithm has been tested on a more physically realistic system, which means that the method has been applied to the problem of translocation of stiff DNA through a pore. The effect of DBF has been measured in the translocation process. The quantitative values of the Potential of Mean Force (PMF), defined as an integral of the effective force on the DNA center of mass depending on its location, is determined through Langevin Dynamics simulations.

The stiff DNA is fixed on the central axis of the pore, in specified set of locations that covers the interval from one side of the pore to the other. With this setup it is possible to determine the force on the DNA center of mass in each given location. The PMF profile was determined from the values of the integral of centre of mass forces.

The measured center of mass forces show that the stiff DNA has to overcome a certain force barrier when entering the pore. This barrier is strikingly higher when dielectrics is taken into account (i.e. when the ICCP^{3}M is switched on). Another barrier is seen at the center of the pore. This implies that the stiff DNA must overcome a significant potential barrier to complete the translocation.

The conclusions presented in this thesis have also appeared in:

"Ensembles Inequivalance in Single Molecule Experiments,
M. Süzen, M. Sega and C. Holm".
Physical Review E **79***, 051118 (2009)*

"A fast method for computing induced charges on the arbitrary inhomogeneous dielectric boundaries in a simulation of large systems,
S. Tyagi, M. Süzen, M. Sega, C. Holm and M. Barbosa."
To be submitted.

Here is the general outline of the chapters in the thesis:

- In Chapter 2, the basic ways of describing biopolymers are reviewed from a statistical and polymer physics point of view. Common SMEs and their applications to nanopores are also introduced as motivation to this thesis.
- In Chapter 3, the theoretical descriptions of quantifying inequivalence of ensembles in polymer stretching experiments will be reviewed and simulation results presented under different settings, after which they will be compared to theoretical results. The analysis of ensemble equivalence in the different force regimes will be given.
- In Chapter 4, a generic and feasible algorithm ICCP^{3}M will be developed for dielectric boundaries in molecular simulations supported by benchmark results on selected geometries when analytical solutions are known, and they will be compares to other methods.
- In Chapter 5, the ICCP^{3}M method will be applied to a specific problem, which is that the effect of dielectric interfaces on the potential of mean force of a coarse-grained DNA model inside a cylindrical pore is studied.
- In Chapter 6, the main results, along with possible future research directions will be discussed and presented.

2 Single Molecule Experiments

This is a general overview of single molecule experiments and takes the focus of our attention from applications to biological physics. Section 2.1 is a historical background of polymers. Section 2.2 is a brief review of basic concepts of polymer physics. In Section 2.3 Single Molecule Experiments (SMEs) and related problems that we address in this thesis are outlined. In particular there is a short description of the elasticity and stretching properties of biopolymers such as DNA. Section 2.4 summarizes some relevant recent developments on SMEs performed with nanopores, and present corresponding theoretical and computational modelling approaches.

2.1 Historical Background

The development of human history and society is defined in terms of tool-constructing technologies available in those particular periods in time. Civilization itself has been influenced by the availability of materials and by the ability to manipulate the materials. In fact, prehistoric societies are broadly divided into Stone Age, the Bronze Age, and the Iron Age labeled after the kinds of tools employed to assist the cause of survival [62]. Similarly, it can be asserted that just as the 20th century can be considered as the Silicon age, perhaps the 21st century is the Polymer age i.e. an age where nano systems in general play a dominant part.

The human body has been using naturally occurring polymers, called biopolymers, for centuries in the process of evolution. Moreover, people have always used polymeric materials such as natural rubber. However, the theoretical understanding of these materials and their physical properties has become possible only after the development of Physics, especially the area of statistical physics. The 20th century has given to us the invention of computers, revolutionary experimental techniques in the manipulation of small objects and related physical theories. All these factors have immensely increased our knowledge of these small systems. The implications are not only important from a practical point of view (such as composites, solar cells etc.) but also crucial in the understanding of the structure and function of macromolecules.

Chemists have started to produce synthetic macromolecules in the middle of the 19th century, but during that period there was no established theory or no experimental evidence of their struc-

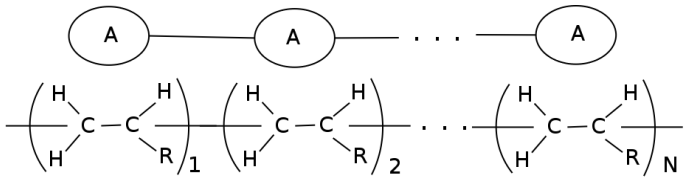


Figure 2.1: Sketch of a linear vinyl polymer of N monomers, where R stands for a generic residue. Adapted from [62].

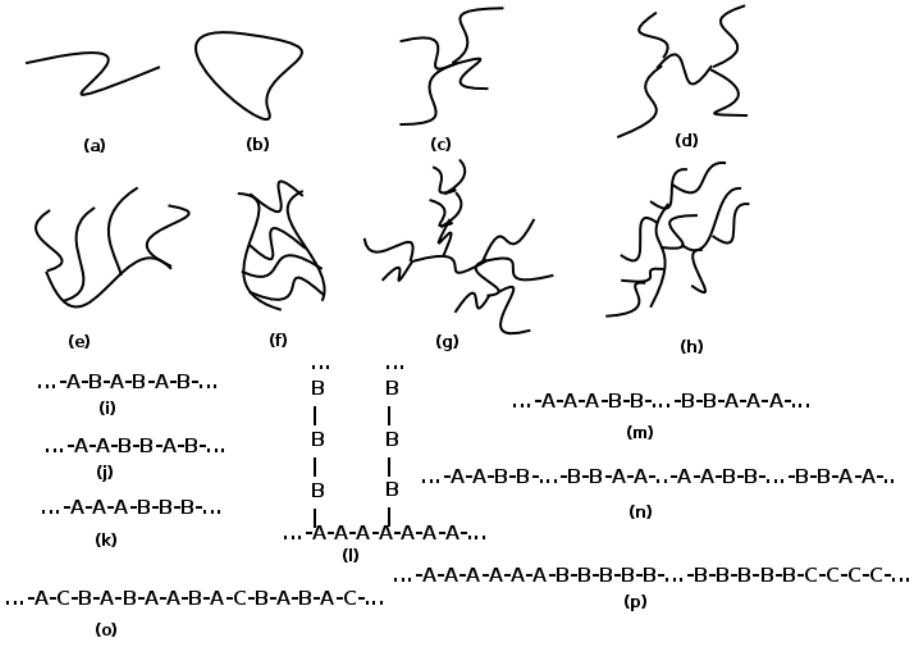


Figure 2.2: Basic classification of macromolecules, homopolymers and heteropolymers (copolymers), based on their monomeric squences and structural architecture, adopted from [62]. Examples of different architectures (a) *linear*, (b) *ring-branched*, (c) *star-branched*, (d) *H-branched*, (e) *comb*, (f) *ladder*, (g) *dendrimer*, (h) *randomly-branched*, copolymers (i) *alternating*, (j) *random*, (k) *diblock*, (l) *graft*, (m) *triblock*, (n) *multiblock*, and terpolymers (o) *multi-random*, (p) *triblock*

ture. Although in 1920 Hermann Staudinger, put forward his hypothises about *polymers* [62], it took some time before their existence was proved [62], especially the work of Ziegler and Natta [63] on the development of experimental techniques in polymerization should be mentioned. The work of many scientists such as Flory and de Gennes contributed to the theoretical foundations of statistical mechanics of Polymers and polymer science in general.

On the other hand different experimental tools were developed in the late 20th century to manipulate small systems and express work on them. As examples, we can cite the Coulter counter for small particle size and concentration detection i.e. through analysis of the conductance signal of the small pores that particles pass through [5], patch-clamp technique for voltage measurements of ion channels [64] and force microscopy technique to measure very small forces in molecular systems [65].

2.2 Polymers: Elementary concepts

The word *polymer* coined by Berzelius derived from the Greek words $\pi o\lambda\nu$ (poly), meaning "many"; and $\mu\epsilon\rho o\sigma$ (meros), meaning "part" [62]. *Linear polymers* are macromolecules consisting of many repeating units, called *monomers*. Monomers are connected to each other by covalent bonds. The number of monomers in a polymer is usually called its *degree of polymerization* N. Hence its contour length can be scaled by $N-1$. Each monomer is considered to be a single entity; therefore we can label a polymer similar to a N-tuple repeating sequence $-A-A-A-...-A-$.

If the sequence contains only one type of monomers, then the polymer is called *homopolymer*. see Figure 2.1.

Many different types of polymer architectures can be constructed depending upon the characteristics of their monomer sequence and network connections. Some basic types of polymer architectures are sketched in Figure 2.2. They are *linear*, topological connections of *ring-branched*, *star-branched*, *H-branched*, *comb*, *ladder*, *dendrimer* and *randomly-branched* polymers. Special attention has been given to randomly-branched polymers due to their wide range of applications.

Macromolecules that contain two different types of monomers are called *copolymers*. Copolymers can be *alternating*, *random*, *diblock*, *graft*, *triblock*, and *multiblock* depending upon their squence and monomeric connections, as depicted in Figure 2.2 (i)-(m). Furthermore, much more complicated polymers i.e. heteropolymers are depicted in Figure 2.2 (n)-(p).

Many biopolymers are heteropolymers. DNA for example, is a heteropolymer with four different types of basic units (nucleotides). All natural proteins are also heteropolymers that generally consist of 20 different types of basic units (amino acids).

The characteristics presented above does not change in the polymerization process, unless covalent bonds are broken, but the *conformation* of a polymer-chain can change. The conformation of a chain is the special structure of a polymer formed by the relative locations of its monomers [62]. The conformation of a chain depends upon three main effects, namely the flexibility of the chain, interactions among monomers, and the surroundings.

Info Box 1 ***Configurational Entropy of a single chain on discrete states*** *[66]*

Let a polymer-chain be defined as a subset of N-sites lattice having 2 and only 2 nearest neighbour connections, except head and tail sites that have 1 and only 1 nearest neighbour connection. The chain length is the number of occupied n-sites on the lattice. The Configurational Entropy $\mathscr{S}$ of this chain in the given lattice sites, with $k_BT = 1.0$, will be

$$\mathscr{S} = ln\mathscr{R}(N, n),$$

where $\mathscr{R}(N, n)$ denotes the number of possible distinct conformation of a chain with length n, see Figure 2.3.

The characteristic flexibility of a chain plays a crucial role because of its physical properties. Some macromolecules are stiff like a wooden stick, while some others are flexible like a snake. The interaction between monomers might be of any character: purely attractive, purely repulsive or a combination on different scales. Chains also interact with the environment (solvent) they are located in. The relative strength of these interactions can change with temperature. The combination of these interactions can drastically change chain conformations and entropy, thus, exerting a crucial influence on the physical properties.

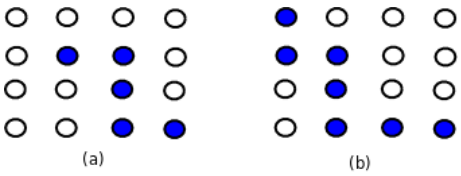


Figure 2.3: Polymers consist of 5 and 7 monomers on the $4x4$ lattice, (a) and (b), respectively. The configurational entropy of a polymer is proportional to number of possible distinct conformations of the polymer in the given configurational space, here the $4x4$ lattice.

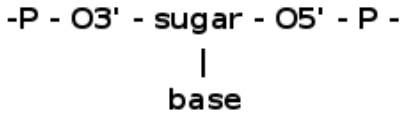


Figure 2.4: Primary structure of macromolecules composed of nucleic acid that are build from the same motif. For example, DNA and RNA differ in their sugar structure.

2.2.1 Biopolymers

A basic motif or the building block (mononucleotide unit) of macromolecules of a nucleic acid is depicted in Figure 2.4. The DNA and RNA contain sugars deoxyribose and ribose, respectively: the acronyms stand for deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA and RNA differ in their sugar structures and nucleic acid bases (nucleotides). Nucleotides may contain several different types of bases such as cytosine (C), thymine (T), uracil (U), adanine (A) and guanine (G). DNA and RNA use the following alphabet (C, T, A, G) and (C, U, A, G), respectively. Specifically for DNA, base pair up as shown in Figure 2.5.

DNA, RNA and proteins are linear heteropolymers. The building blocks of proteins are peptides to which 20 different side groups can be bound, forming 20 different amino acids [10]. Double stranded (ds) DNA has a double-helical structure, independent of its sequence. On the contrary, the structure of proteins and some RNAs (transfer: tRNA, ribosomal: rRNA) is uniquely determined by their sequence. Because of this, one of the central goals in the theoretical biopolymer research is to infer the structure of a polymer from the knowledge of its sequence.

Biopolymer structures have an immense length and number of sequences in comparison to the

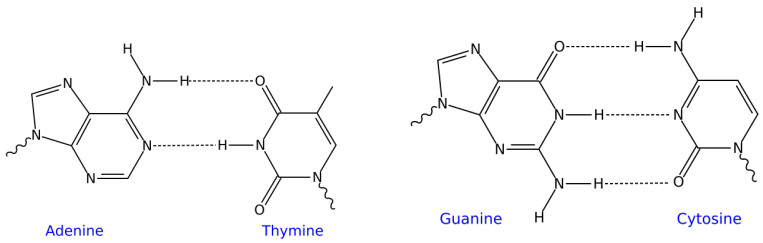


Figure 2.5: Chemical Structure of DNA base pairs: Adanine/Tymine and Guanine/Cytosine.

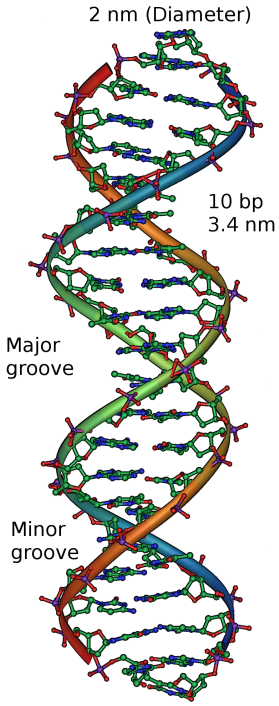


Figure 2.6: Right handed DNA double-helix. An alphabet of C,T,A,G is used to form the base pairs of 2 nm diameter. A major and minor groove of 22 and 12 Å is shown.

size of the biological organism. For example, a typical DNA sequence contains $10^7 - 10^{10}$ units, its typical length of 2 m can be compactified in a densely packed genome that would form a ball of just 2 μm diameter due to huge aspect ratio of contour length versus diameter (2 nm). Hence, the DNA fits into the nucleus but a suitable compaction mechanism is required [67].

2.2.2 Hierarchical Structures

Biopolymers possess a hierarchy of structures that are called; primary, secondary, tertiary and sometimes, quaternary. The simple sequence of repeating units in the chain is denoted as the primary structure. The positions of monomers determine the secondary and tertiary structures in the short-scale and long-scale order respectively [10]. The arrangement of many folded protein molecules in a multi-subunit complex is called quaternary structure.

Numerous environmental factors (the solvent, the ionic conditions and the temperature) affect the structural properties of double stranded DNA, sketched in Figure 2.6, such as the effective diameter and helical pitch. For example, B-DNA [68] is the canonical form of DNA that is embeded in an aqueous environment, forming a right-handed double helix. B-DNA has the following parameters: a diameter of about 24 Å, the helical pitch of 10.4 base pairs per turn and bases separated vertically by roughly 3.4 Å. The double-helix is not perfectly symmetrical, but is instead characterized, see Figure 2.6, by a major groove and minor groove, which have sizes of about 22 Å and 12 Å, respectively.

The path of the double helix central axis determines the tertiary structure. DNA can have more

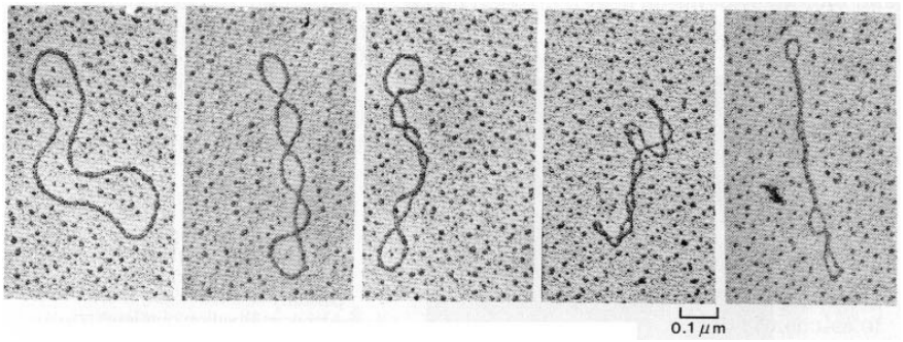


Figure 2.7: Tertiary structures: Images of supercoiled plasmids (5 kb) obtained by electron microscopy [69]. Degree of supercoiling is increasing from left to right. Images show how supercoiling compacts DNA.

compact tertiary structures: super-coiled plasmids, observed by electron microscopy [69], exhibit compact structures known as 'plectonemes', reminiscent of the intertwining observed on an entangled phone cord shown in Figure 2.7.

2.3 Experimental Techniques

Biological processes, once inaccessible, can now be investigated quantitatively by means of single-molecule manipulation techniques that revolutionized our understanding of biological physics and small systems in general [1, 3]. Techniques such as optical tweezers [70, 71, 72, 73, 74, 75], atomic force microscopy [65, 76, 77, 78, 8, 79], magnetic tweezers [80, 81, 82], elongational flow [83, 84, 85] and other force microscopies [1, 9, 10, 2] have enabled researchers to perform quantitative measurements of small forces and allow us to study a large spectrum of individual molecular systems.

The primary information extracted from an atomic force microscope, for example, is usually presented as force-extension curves, a force applied to the molecule versus its extension [2]. The force-extension relations or, in thermodynamic terms, the mechanical equation of state, can be measured and calculated under different protocols: (i) It is possible to fix the extension of the macromolecule and measure the resulting force necessary to maintain this extension (isometric protocol). (ii) It is possible to apply a given force and measure the resultant extension of the molecule [24, 25] (isotensional protocol). It is quite important to formulate the right protocol for the stretching of macromolecule in an SME. The realization of these two protocols in experiments are made possible by adjusting the force constant of traps, which in the case of AFM trap is manifested by the cantilever. So far stretching experiments were done (approximately) under (ii), mainly for reasons of sensitivity [24, 2]. These two protocols in depth and their equivalence conditions will be investigated in depth by means of computer simulations in this thesis, in Chapter 3. Normally we take it for granded that, statistical mechanics tells us that the outcome of a measurement does not depend upon the used measurement protocols in the thermodynamic limit [86]. This assertion will be studied in polymer stretching experiments for a range of forces by employing Langevin Dynamics simulations to clarify the recent debate on the validity of this assertion for single polymers [17, 30].

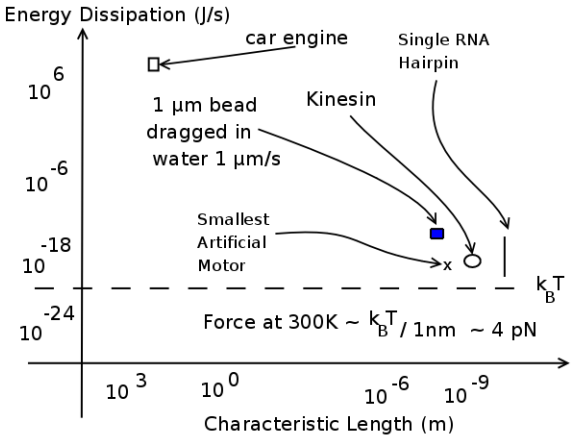


Figure 2.8: The length-energy scales for engines of life. The energy dissipation rate of a car engine is almost 20 order of magnitude higher than the rate of the biological motor. The overall picture of length scales would help us visualize at what level our experimental systems operate. Adapted from [12].

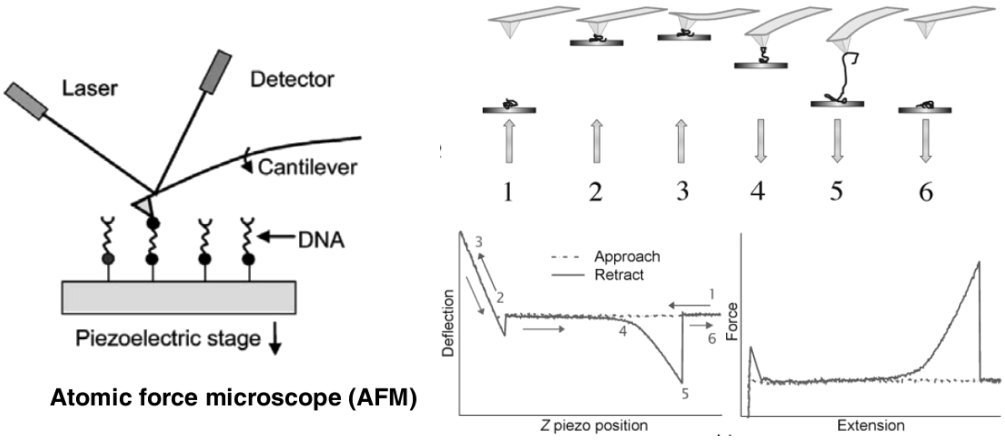
The measurements of the ionic current change in a single ion channel, for example, in cellular membranes [64, 5] or solid-state nanopores [6, 7] immensely increase the understanding of transport phenomena in this context. However, constructing computer simulation models of these systems give rise to a computational challenges, such as determination of electrostatic forces on the ions or polymers due to dielectric interface, formed by electrostatic continuum domains, solvent and the membrane. A new solution to this algorithmic problem has been proposed. The mathematical description of the problem, previous proposed solutions, a novel method, and the implementations, tests and comparisons will be addressed in depth in Chapter 4. Moreover the effect of dielectric interfaces on DNA translocation through a channel will be examined at the end of the thesis.

2.3.1 Forces and Energies in the Single Molecule Scale

The range of forces, energies and length scales that are relevant to SMEs, must be outlined. Small systems are studied differently from their macroscopic counterparts, since thermal fluctuations with typical energy $1\ k_BT = 4 \cdot 10^{-21} J = 0.6$ kcal $\cdot mol^{-1}$ at $T \sim 300K$ cannot be neglected. In this context, we compare dissipation rates of typical engines in different length scales in Figure 2.8. It can be seen that the energy scales dealt in biological motors are almost 24 orders of magnitude smaller than the conventional combustion engine. The question is whether macroscopic thermodynamics would work on this scale. This question is addressed in Chapter 3 investigating the equivalence of thermodynamic equations of state of polymer stretching experiments via computer simulations.

The theoretical challenges in describing nanoscale engines occur due to the noisy environment and out of equilibrium conditions in case of SMEs. For example, typical forces that occur on the nanometer scale are of the order of $1\ k_BT/nm = 4pN = (4 \cdot 10^{-12} N)$; for a cell this can be as much as $10\ pN$ and it is comparable to its own weight.

The work should be done against the configurational entropy of the system consisting of the



(a) Sketch of a typical AFM setup [2]. An AFM tip sticks to a macromolecule's one end, while its other end is fixed on the stage. One can measure the deflection of cantilever via the detector.

(b) The stages of the AFM stretching procedure [79] from capture to release of a macromolecule, labeled from 1 to 6. The corresponding stages of the piezoelectric position versus the deflection of the macromolecule; the force on the end of the macromolecule versus its extension curves are shown.

Figure 2.9: Sketch of a typical AFM setup.

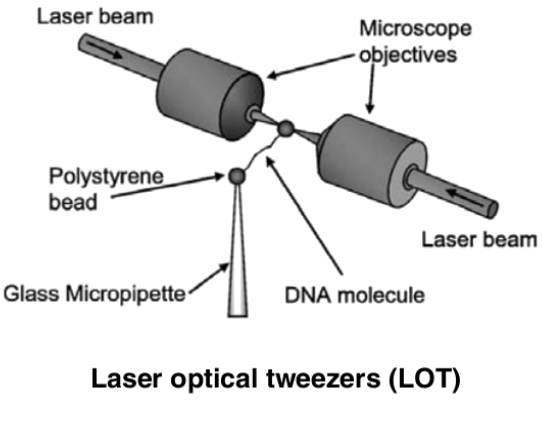
molecule (e.g. protein, DNA) and its solvent (water, ions), to bring the molecule in fully stretched configuration with SMEs [66]. In this type of stretching, the basic unit of energies is expressed in bond energies i.e. in electron volt $1\ eV = 1.6 \cdot 10^{-19} J = 36\ kcal/mol$. Hence, the forces necessary to break receptor/ligand bonds are in the order of $1 eV nm^{-1} = 160\ pN$ [10].

2.3.2 Atomic Force Microscopy

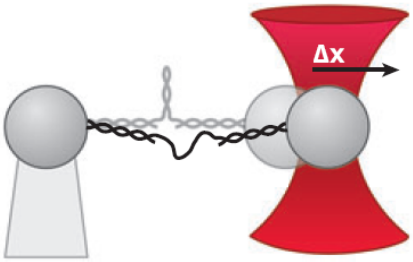
Perhaps one of the best known SME techniques is based on the usage of an Atomic Force Microscopy (AFM), a descendant of the scanning tunnelling microscopy applied to obtain Ångström resolution images of surfaces [65].

The basic principle behind an AFM is that a cantilever (with a given stiffness, $10 - 1000$ pN$\cdot nm^{-1}$) with a tip can stick (or be functionalized to stick) to a monomer of a biopolymer. The cantilever can exert a precise force on a polymer. The polymer extension can be measured via the observed deflection of the cantilever. This set-up is depicted in Figure 2.9 (a) [1, 2, 79]. The conventional experimental measurement procedure is summarized in Figure 2.9 (b): Initially the macromolecule is captured (label 1-3), pulled (label 4), its extension measured (label 5) and finally the macromolecule is released (label 6).

There are many applications of AFMs, including surface scanning and detection. The use of the AFM for biomolecular systems [76, 77, 78, 8, 1, 79] specifies the stretching or force-extension relation of the biopolymer. The AFM covers forces in the 20 pN-10 nN range depending on the stiffness of the cantilever. Even though the AFM is quite a powerful tool in manipulating single molecules, unwanted interactions (van der Waals, electrostatic and adhesion forces) between tip and substrate may take place, and reduce the accuracy of measurements.



(a) One end of the DNA molecule is anchored while the other end is attached to a bead that is under the influence of a optical trap due to the potential generated by a Laser beam [2].



(b) The trapped bead can be pulled by the laser tweezer, hence, the anchored DNA can be stretched, and the extension can be measured by inferring the Δx via the value of the harmonic trap force [75].

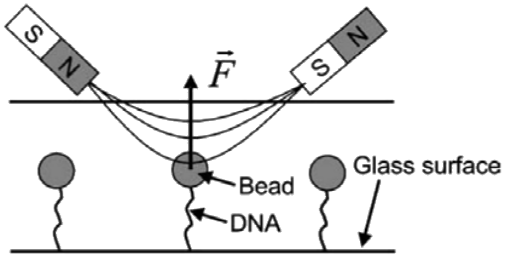
Figure 2.10: Sketch of a typical laser optical tweezers set-up.

2.3.3 Laser Optical Tweezers

A particular and very useful experimental technique is to use laser optical tweezers to measure forces on the nanoscale. The technique is based on the principle that light can exert forces on particles [71, 70, 87]. A typical laser optical tweezer is depicted in Figure 2.10. The force exerted by the laser on the trap can vary within the range $0.1-100$ pN and depends upon the bead size and the laser power. A typical bead size is of order of $1-3\mu m$. The laser power should not exceed 100 mW to prevent heating of the bead. In general, the forces acting on the bead are approximated by a harmonic trapping potential. The force is given by $F = k\Delta x$, where k is the stiffness constant of the trap and Δx is the distance between the bead and the centre of the trap, see Figure 2.10. The force resolution of laser optical tweezers is usually of the order of 0.1 pN [2]. This value is at least 10 times more accurate than in AFM, because the stiffness of the trap is 100 times less than that of a cantilever. Therefore the force resolution makes optical tweezers an ideal tool to investigate the behaviour of biomolecules [72, 73, 74, 75]. A disadvantage of this technique is that one cannot apply as large forces as with an AFM.

2.3.4 Magnetic Tweezers

The principle that a magnetized bead feels a force when it is embedded in a magnetic field gradient $F = -\mu\nabla B$ lies at the heart of the design of magnetic tweezers (MT). The basic setup is shown in Figure 2.11. The trapping of magnetic bead is provided by the magnetic field generated by two magnets. The typical force range that can be measured with this technique is $10^{-2} - 10pN$, which is highly sensitive to the size of the magnetic bead. Due to the low stiffness of this trapping potential of the magnetic trap, which has typically a stiffness of about $10^{-4}\ pNnm^{-1}$, one can



Magnetic tweezers (MT)

Figure 2.11: Sketch of a typical magnetic tweezers set-up [2]. One end of the DNA molecule is anchored while the other end is attached to a bead that is under the influence of magnetic trap due to the potential generated by magnets.

measure very low forces.

The forces which can be measured using this technique are one million times smaller than in AFMs and a thousand times smaller than in optical tweezers. The MTs have been extensively used to investigate elastic and torsional properties of DNA molecules [80, 81, 82]. However, a disadvantage of MT is that one cannot apply as large forces as in AFM or optical tweezers; hence, its range of applicability is limited.

2.3.5 Stretching DNA: Flexibility

The general term *flexibility* embraces both static aspects (curvature) and dynamic aspects related to the existence of a double helix departing from the ideal linear rigid model [88]. The conformational studies of DNA and its double-helix structure show that there is only a limited range of angles i.e. bond angles that characterize the biomolecular chain in the nucleotide chain: in other words, there is some rigidity in the structure. DNA to begin with should be described as a fairly rigid molecule, at least over small lengths of the chain. However, the real problem is in fact that the DNA chains fold in multiple ways.

The study of flexibility must therefore begin with an overall look at the whole DNA molecule, which involves the general principles of polymer mechanics in a given medium. A more detailed analysis on a local scale will then be needed in order to take into account the squence and any irregularities in the geometry of the double helix.

As any polymer in a good solvent, the DNA in an aqueous buffer adopts a random coil conformation that maximizes its entropy [66]. When stretching the molecule one encounters at low forces ($F < 10pN$) an entropy dominated regime and then up to about 70 pN an elastic Hookean regime where the DNA stretches like a spring. Beyond that regime, the DNA undergoes structural transitions.

The first measurements of the entropic elasticity of a single DNA molecule of crystallographic length of 16.2 μm were reported by Smith et al [80]. They used a combination of magnetic fields and hydrodynamic drag to pull with a force F on small super paramagnetic beads tethered to a surface by a single DNA molecule. The primary information extracted from this type of

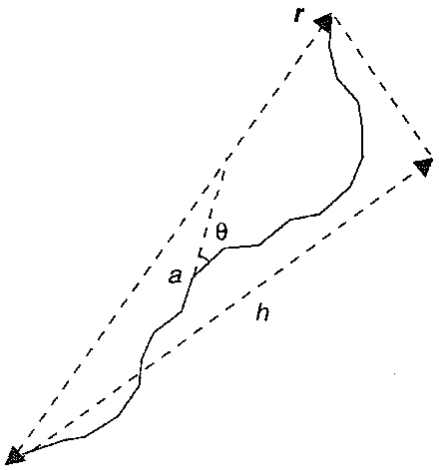


Figure 2.12: Planar model of Kratky and Porod (worm-like chains) consisting of N segments of length a, making an angle θ with adjacent segments. In this example, h is the projection of the vector $\mathbf{r}$ joining the two ends to the first segment [88].

experiments is a *force-extension curve*. The precision of their measurements has been good enough to invalidate the theoretical prediction of the freely jointed chain (FJC) model that treats the DNA as a chain of freely jointed rods whose entropy is reduced upon stretching. A physically more sensible approach that considers the DNA to be a continuous flexible worm-like chain (WLC) has later been solved by Marko and Siggia [18]. It computes the entropy change of the chain more realistically than the FJC model, particularly, at high forces where the entropy is underestimated in the FJC model (which neglects bending fluctuations of the discretized rods) [10].

The worm-like chain (continuous curvature)

A polymer model has been proposed by Kratky and Porod (WLC) [89, 10, 88] to describe all states between two extreme models of the perfectly flexible chain with free rotation (or Gaussian chain) and the prefectly rigid rod-shaped chain. The wormlike chain is a mathematical minimal model that captures universal aspects of the statistical mechanics and dynamics of stiff and semi-flexible polymers.

The WLC depicted in Figure 2.12 consists of a freely joint chain (FJC) of N segments, each of length a, of a perfectly flexible chain [88] with the segments unrestricted in their respective orientations [10]. We seek the probability distribution function $p(\mathbf{r})$ for the FJCs end-to-end vector. Mathematically, this situation is equivalent to random walks in three dimensions. Let $\mathbf{r}_1, ..., \mathbf{r}_N$ be the individual independent steps or displacements, all subject to the same probability distribution function $p_i(\mathbf{r}_j)$. Let $\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2 + ... + \mathbf{r}_n$ be the resulting end-to-end displacement after N steps with probability distribution $p(\mathbf{r})$. For large N, the probability distribution for the freely joined

chain in three dimensions approaches a three dimensional normal or Gaussian distribution:

$$p(\mathbf{r}) = \left(\frac{3}{2\pi N a^2}\right)^{3/2} \exp\left(-3\mathbf{r}^2/(2Na^2)\right).$$

How to derive this is shown in Chapter 3 where the FJC model is used in our Langevin Dynamics simulations for investigating the ensemble equivalence of stretched chains. But here the flexibility of a chain namely its *persistence length* is characterized as follows [88]: Consider N segments of length a each making a small angle θ with the previous one, lying on a cone of vertical semi-angle θ around the previous segment, see Figure 2.12. The mean value $\langle h \rangle$ of the projection on the first segment of the end-to-end distance is given by

$$\langle h \rangle = a \sum_{k=0}^{k=N} x^k = a(1 - x^N)/(1 - x),$$

where $x = cos\theta$. The *persistence length* L_p is defined as the limiting value of h as $N \to \infty$. We then have:

$$L_p = a/(1 - x),$$

(As $\theta \to 0$, the chain can be viewed as one for which L_p remains finite, i.e. in which $a \to 0$, which amounts to introducing a continuous curvature). Since θ is small, $\cos\theta \approx 1 - \theta^2/2$ and

$$L_p = 2a/\theta^2.$$

Note that the persistence length does not depend on the length L along the curve i.e. Na, but is an intrinsic property of the polymer in a given medium. It is related to the structure and the interactions and not to the molecular mass. We can replace $x^N = (1 - \theta^2/2)^N$ by $\exp(-N\theta^2/2)$ and putting the total length of the chain $L = Na$. We obtain

$$x^N = \exp(-L/L_p)$$

and hence

$$\langle h \rangle = L_p \left[1 - exp(-L/L_p)\right].$$

As $L \to \infty$, $\langle h \rangle \to L_p$, which agrees with the definition. If, on the other hand, the chain is small enough for $L \ll L_p$, we obtain $\langle h \rangle \approx L$. This short chain behaves like a rigid rod of length L. With this *worm-like* model, we can thus account for both local rigidity and the flexibility of a sufficient long chain.

There has been a recent debate on whether the outcome of measuring force-extension curves depend on the chosed experimental conditions, i.e. is there an ensemble difference, and if yes, is this difference vanishing [17] or not [30] in the thermodynamic limit (very long polymer) in weak force regimes. To clarify this debate, completely flexible chains have been used, the FJC model in polymer stretching Langevin Dynamics simulations in Chapter 3 producing two different data sets for force-extension curves of different polymer length $L = Na$. The first data set is the isotensional protocol and the second is the isometric protocol. Even though this model is quite simple it enables one to capture the essential physics behind the ensemble equivalence relation of the two protocols in performing polymer stretching experiments, where equivalence of protocols implies obtaining the same outcome. It will be clarified in this thesis under which conditions these protocols are equivalent by employing Langevin Dynamics simulations for producing force-extension curves in two different ensembles.

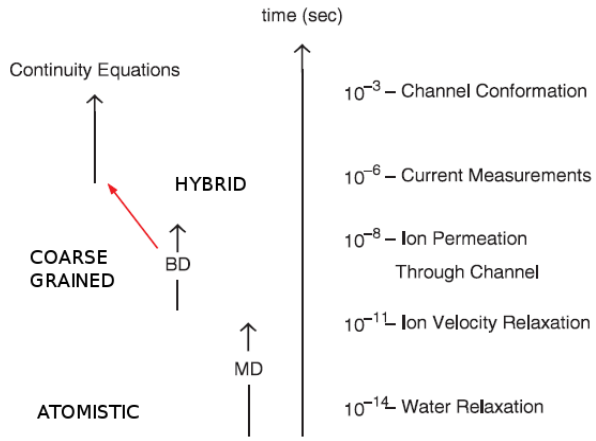


Figure 2.13: Time scales of channel processes and suitable modelling techniques [93]. Simulation of atomistic scale dynamics can be achived with Molecular Dynamics (MD), but higher the system size coarse graining is inevitable, such as Brownian Dynamics (BD) or Langevin Molecular Dynamics. In the mesoscale hybrid methods can be employed.

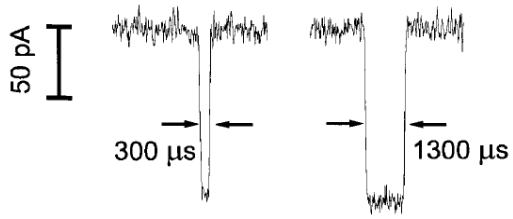


Figure 2.14: Time evolution of current measurements. Two typical blockades with lifetimes of 300 and 1300 ms due to translocation of a biopolymer through a biological pore [94].

2.4 Nanopores

One of the most important processes of living organism is the material transport (ions or biopolymers) through membrane channels or nuclear pores [5, 4, 90]. This machinery is mimicked in technological applications such as micro/nanofluidics devices [7] and nanometer sized pores (nanopores). Theoretical and computational modeling of these nanoscale systems have taken place in modern research in physics, chemistry and biology [91, 92, 5, 7].

2.4.1 Molecular Sensing with Biological Pores

A wide range of biological functions are controlled directly by ionic channels [4]. They control the transport and mobility of ions, hence the electrical properties across the membranes that are vital for functions of biological cells [95, 96]. Moreover, ion channels control electrical signaling in the nervous system, coordinate the contraction of muscles, regulate uptake of material into the biological cell and a lot more, having complex properties and functioning as nanodevices [90]. The detection of single-channel currents is possible by the patch-clamp technique [64]. This

opens a large window of opportunities for ion-channel research [97, 98] as well as theoretical and computational approaches [91, 98, 92].

Selectivity, permeation, charge character and the correlations between structure and function can be counted as the main investigation themes in ion channel research [99]. The dominant effect of dielectrics in the modeling of ion channel dynamics has been pointed out frequently in the literature [42, 100].

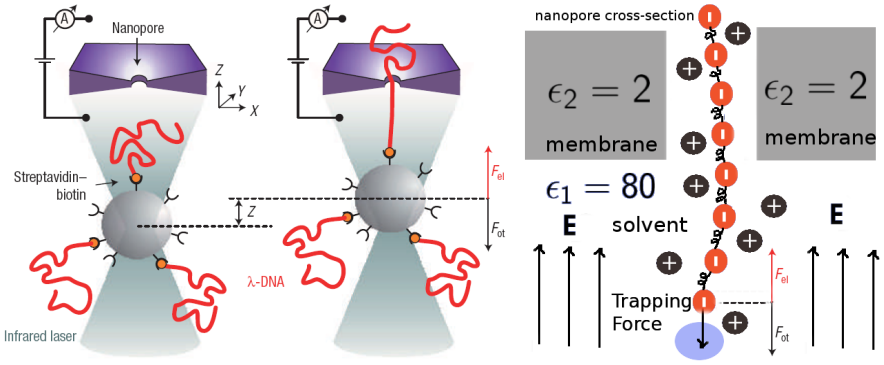
Due to the inherent complexity and large spatio-temporal scales involved in these systems, computational modelling of ionic channels and related process are quite challenging. A typical time scale of the processes and the possible corresponding modeling techniques are summarized in Figure 2.13. If the time scales of solvent and membrane degrees of freedom is much smaller then the translocation time scale of the macromolecule through the nanopore, these degrees of freedom can be coarse-grained, using a dielectric continuum (DC) representation. Similarly for the macromolecule, all chemical details may not be needed; hence a reduced representation would be sufficient to capture the essential physics. A computational problem occurs within this approach, namely how to compute forces generated by dielectric boundary effects efficiently and accurately. This problem is addressed in depth in Chapter 4 and a novel solution is proposed.

Apart from the patch-clamp technique, this is another well known technique, called Coulter counter, which works by forcing small particles with a pressure difference to flow through a small hole causing blockades in the ionic current [5] so that its size and concentration can be measured directly by analysis of the conductance of the channel. The same principle has been employed for investigating biopolymer translocation through a biological nanopore driven by electrical voltage [94, 101, 102]. Typical current characteristics during current blockade are shown in Figure 2.14. Current blockades are sudden drops in the channel current due to the presence of a macromolecule inside the channel. This current drop appears due to the fact that when a molecule is inside the pore, it prevents permeation of most of the ions. These polymer translocation experiments and their findings have been analyzed in detail both theoretically [103, 104, 105, 106, 107, 108, 109, 110, 111] and computationally [112, 113, 114, 115, 116, 117, 118], and continue to attract a considerable research interest, because of its biotechnological value. One of the hopes is that possibly the sequence of a macromolecule can be extracted from current characteristics. Almost all of the simulation studies in the literature try to generate scaling laws for translocating polymers, such as polymer length versus translocation time. However, up to now there is almost no work that addresses the effect of the dielectric boundary force on polymer translocation. For this reason, in this thesis the effect of dielectric boundaries on the translocation of short DNA through a nanopore is investigated in Chapter 5.

2.4.2 Molecular Sensing with Solid-State Pores

All experiments discussed so far for molecular sensing are restricted to naturally occurring protein membranes; however, advances in nanotechnology now allow the fabrication of artificial solid-state nanopores as insulating membranes [6]. Both biopolymer translocation [119, 120, 31] and ionic tranport problems [121, 122] are addressed using solid-state membranes. Hence, the solid-state nanopores are versatile new single-molecule tools for biophysics and biotechnology [6].

As a show case example, we discuss a novel technique that combines an SME, optical tweezer and a nanopore in measuring force on voltage driven DNA entering a solid-state nanopore through ionic-current detection [31]. This novel method directly measures the force acting on the chain, while a part of it is inside the channel as a function of the driving field with different electrolyte



(a) Experimental sketch of a DNA molecule inserted into and pulled out from a nanopore by optical tweezers [31]. A force exerted on the end of the DNA when it is partially inside the nanopore can be measured by this technique. Variation of this measured force against the external electric field leads to a direct measurement of the effective charge of the DNA.

(b) The sketch of a coarse-grained model of the given experimental situation. Solid-state nanopore and solvent modeled on dielectric continuum, with typical dielectric constants of 80 and 2 respectively. DNA is represented as a bead-spring model. Counter ions are placed to keep the system electroneutral, where they physically represent associated ions present in the system.

Figure 2.15: Experimental setup of pulling DNA out of a nanopore and corresponding coarse-grained model.

concentrations. The scheme of the experiment is sketched in Figure 2.15 (a). Here the opposing force exerted by the optical tweezers slows down and even arrests the translocation. For the first time, it turns out to be possible to determine the charge of DNA directly. The value of the force was $0.24 \pm 0.02 pNmV^{-1}$ that corresponds to 0.50 ± 0.05 $electrons/basepair$ independent of salt concentration [31]. A coarse-grained model that is convenient for Langevin Dynamics corresponds to the experimental situation depicted in Figure 2.15 (b). The problem of handling the electric field computation efficiently in the presence of dielectric interface that appears due to membrane and solvent dielectric continuum domains as shown in Figure 2.15 (b) is addressed in depth in Chapter 4. A novel and fast computation method has been developed for the forces that occur due to the dielectric interface of realistic simulation of the experimental situation like the one sketched in Figure 2.15 (a). A coarse-grained bead-spring like DNA model will be presented along with its translocation through a nanopore in Chapter 5, where a nanopore geometry similar to the one depicted here is employed.

2.5 Summary

A short introduction to the fundamental notions of polymer physics is given from the biophysics and statistical mechanics point of view. The basic working principles, applicability range and the importance of SMEs in cellular biophysics and soft condensed matter physics have been described, focusing only on a small fraction of SMEs: AFM, Optical and Magnetic Tweezers. Special at-

tention has been paid in investigating force-extension characteristics of macromolecules. The question of possible difference in the force-extension curves of different realizations of polymer stretching experiments, (i) constant extension, and (ii) constant force protocols, rised as a fundamental question in force microscopy techniques.

A recent debate on the whether these two protocols are equivalent will be clarified in this thesis by empolying large number of Langevin Dynamics simulations of polymer stretching. A brief introduction to worm like chain (WLC) model convenient for polymer stretching simulations is outlined in this chapter.

The growing body of experimental and theoretical efforts devoted to study ionic and biopolymer transport through channels, describing their apparent importance in sequencing is briefly discussed. Different modelling strategies are developed due to the varying complexity of these systems. One of the most important computational challenges in describing the membrane-solute as a dielectric continuum domain, namely: computing forces due dielectric boundary is emphasized. An example mapping from a experimental situation that employs combined SME and a nanopore to coarse-grained simulation model is shown and explained in Figure 2.15.

In this thesis, an efficient and accurate algorithm that can compute forces due to dielectric boundaries has been developed. Moreover, the effect of dielectric boundaries on short DNA translocation through a nanopore is investigated in Chapter 5. The main motivation behind these themes is summarized in this chapter.

3 Nonequivalence of Ensembles for Finite Polymers

The present chapter addresses the problem of ensemble equivalence of Gaussian chains from the point of view of computer simulations, employing coarse-grained Langevin dynamics simulations. In Section 3.1, we discuss the basic theoretical background and the appropriate way of quantifying the ensemble equivalence. After we describe the employed simulation techniques in Section 3.2, we present the analysis of the force-extension curves in the Helmholtz and Gibbs ensembles for Gaussian chains with zero and non-zero average bond vector. We investigate the scaling behavior of the ensemble difference with respect to the chain length, and confirm the prediction of Neumann[30] on the ensemble inequivalence in the limit of vanishing applied forces.

3.1 Ensemble Equivalence in a SME

Among the technical developments that influenced the research in molecular biophysics during the last two decades, the possibility of manipulating single molecules by means of different techniques has certainly played a pivotal role, and has been proved to be an invaluable tool to gain insight into the structural properties and function of macromolecules involved in many biological processes [2]. The techniques devised to manipulate single molecules are collectively known as single molecule experiments (SMEs) [9, 10, 2] and include, among others, approaches based on atomic force microscopy [65, 76, 77, 78, 8], optical tweezers [70, 71, 1, 72, 73, 74, 75], fluorescence detection of Förster resonance energy transfer [123], elongational flow [83, 84, 85] and magnetic tweezers [80, 81, 82].

It is widely known that equilibrium thermodynamics can be recovered from a statistical mechanics description in the limit of infinite system size, (the thermodynamic limit) and that, except when in proximity of a phase transition, all statistical ensembles provide the same average values for the observables of interest. In other words, in the thermodynamic limit different ensembles become equivalent.

In case of SMEs, on the contrary, the outcome of a measurement explicitly depends on the control parameters, that is, on the choice of which quantities are kept constant and which ones are allowed to vary. For this reason, the efforts in providing satisfying theoretical descriptions of small, out of equilibrium systems have been intensified during the last decade. In particular, the validation of the hypothesis made by Flory [124], that ensemble equivalence for a SME on a linear polymer should be obtained in the limit of infinite chain length, has been often subject of investigations [2, 17, 30]. Since one of the basic means of extracting information regarding a linear polymer in a SME is to analyze the molecular force–extension curve (FEC) [2], it is natural to introduce two different conjugate ensembles, namely, the Helmholtz, or isometric ensemble, and the Gibbs, or isotensional ensemble. In the isometric ensemble the position of the chain ends is employed as a control parameter, fixing also the end-to-end distance, while in the isotensional case the control parameter is represented by the force applied on one loose end. The conditions

under which the equivalence between these two ensembles can be obtained have been investigated in many works, both from a theoretical point of view [125, 126, 127, 128, 28, 29, 129, 130, 24, 131, 132, 17, 22, 19, 20, 133, 23] and by means of computer simulations [134, 135, 136, 137, 138, 139, 140, 26, 21, 141]. Some of the authors concluded that ensembles are equivalent in the infinitely long chain limit for a Gaussian chain [130, 140], as well as for a generic chain [17] and toy lattice chain models [17, 19, 20]. Some other works stated that ensembles are not equivalent in the thermodynamic limit for a single chain [28, 137, 138, 29]. Since there was no general concensus, we try to settle this question with carefully designed set of simulations.

3.1.1 Concept of Control Parameters

In statistical mechanics the connection with thermodynamics is realized by defining the thermodynamic potentials from the partition function, in the limit that every extensive control parameter is going to infinity. As a general result, given a statistical ensemble and a control parameter, it is possible to construct a conjugate ensemble using as a conjugate control parameter the derivative of the ensemble's thermodynamic potential. It is generally assumed that in the thermodynamic limit, two conjugate ensembles should be equivalent, namely, they should provide the same expectation value for the thermodynamic quantities [86]. Let us take as an example the case of the canonical ensemble, whose partition function will be denoted by $Q(N, V, T)$, and whose thermodynamic potential is the Helmholtz energy A, defined as

$$\beta A = \lim_{(N,V)\to\infty} \ln Q(N, V, T).$$

Here $\beta = 1/k_B T$, where k_B is Boltzmann's constant. The main thermodynamic quantities can then be derived by computing the derivatives of the thermodynamic potentials with respect to their parameters. In the example of the canonical ensemble, by choosing the pressure $p = -\partial A/\partial V$ as the conjugate control parameter, the isothermal-isobaric ensemble (NpT) is generated as a conjugate ensemble. The Legendre transform applied to the Helmholtz energy yields the thermodynamic potential of the isothermal-isobaric ensemble, namely, the Gibbs energy $G(N, p, T) = A + pV$. The conjugate partition function, $\mathscr{Z}(N, p, T)$ can be written in a natural way as a weighted sum of the canonical partition functions

$$\mathscr{Z}(N, p, T) = \int dV Q(N, V, T) \exp(-\beta p V). \tag{3.1}$$

Note that the integral transform given in Equation (3.1) is nothing but the *Laplace Transform* of $Q(N, V, T)$ to function $\mathscr{Z}(N, p, T)$, implying a transform of V dependent function to p dependent one. Consequently, the ensemble average $\langle \mathscr{A} \rangle$ of any observable $\mathscr{A}(V)$ can be expressed in the conjugate ensemble as

$$\langle \mathscr{A} \rangle = \mathscr{Z}^{-1} \int dV Q(N, V, T) \exp(-\beta p V) \mathscr{A}(V). \tag{3.2}$$

The ensemble equivalence problem can then be stated as follow: can the thermodynamic potential of the conjugated ensemble be obtained from the thermodynamic limit of the conjugate partition function? If this is true, the thermodynamic quantities computed in either ensembles will lead to the same expectation values.

Info Box 2 ***Equivalence of Canonical and Microcanonical Ensembles***
[142, 86, 143]

Despite the fact that, the conceptual and operational differences between two ensembles exits, however the principle of equivalence of ensembles guarantees that an observable, for example the internal energy E, can be reproduced, regardless of which ensemble used in the measurement in the thermodynamic limit. Using the microcanonical ensemble, which is characterized by E,V and N, internal energy, volume and number or indistinguishable particles, respectively, the procedure to find $E = E(N, V, T)$ is outlined as follows of the given temperature T.
The Microcanonical Ensemble: Procedure A

1. *Find the number of accessible microstates $\omega(N, V, E)$ (for example all possible configurations of a polymer in the given state-space).*
2. *Calculate the thermodynamic entropy function $S(N, V, E) = -k_B ln(\omega(N, V, E))$.*
3. *Calculate the temperature through the definition $T(N, V, E)^{-1} = (\partial S/\partial E)_{N,V}$.*
4. *Solve th expression for E to produce the desired function $E(N, V, T)$.*

In contrast to the microcanonical ensemble, the canonical ensemble is characterized by T,V and N. Therefore, the procedure to find $E(N, V, T)$ differs. One can find $E(N, V, T)$ via the following steps.
The Canonical Ensemble: Procedure B

1. *Find the partition function (or sum) $Q(N, V, T) = \sum_{\Gamma Microstates} \exp(-E(\Gamma)/k_B T)$.*
2. *Calculate the thermodynamics free energy function*

$$F(N, V, T) = -k_B T ln Q(N, V, T)$$

3. *Use the Gibbs-Helmholtz relation*

$$E(N, V, T) = (\partial(F/T)/\partial(1/T))_{V,N} = -(\partial ln Q/\partial\beta)_{V,T}$$

Equivalence of ensembles implies procedures A and B yield the same energy.

Note*: The equivalence of procedures here will not work if the thermodynamic function ($F(N, V, T)$ or $S(N, V, E)$) is not a concave function everywhere [144]. A concave function can not be express as Legendre transform, hence the energies generated by two ensembles are not equivalent.*

3.1.2 Ensembles for a Single Chain

In experiments in which single molecules are directly manipulated, as is the case for AFM or optical tweezers, the device is basically acting on the terminal part of the molecule, and depending on the strength of the interaction of the tool with the molecule, the former can be employed either to hold firmly one molecular end or, to measure the force which is acting on it, while the other end is attached to some rigid support [17]. Both the isometric and the isotensional ensembles can be realized by changing the force constant of the cantilever in the case of AFM or spring constants in the case of optical tweezers experiments [25], where the use of high force constants will lead to a sampling of the isometric ensemble. It is therefore natural to identify as interesting control parameters the molecular end-to-end (displacement) vector $\mathbf{X}$ and its conjugate variable, the force $\mathbf{F}$ acting on one end. This leads directly to the definition of two corresponding conjugate ensembles, namely the isotensional one, in which a constant force is applied on one molecular ending and, the isometric one, in which the end-to-end vector is kept fixed (note that keeping the end-to-end *distance* fixed, a different ensemble is generated [30]). A schematic representation of the formal analogy between the isometric, isotensional, canonical and isothermal-isobaric ensembles is depicted in Figure 3.1. In this framework, two natural observables emerge, namely, the projections X and F of the control parameters $\mathbf{X}$ and $\mathbf{F}$ onto the end-to-end vector, which read

$$X \equiv |\mathbf{X}| \tag{3.3}$$

$$F \equiv \mathbf{F}\cdot\mathbf{X}/X, \tag{3.4}$$

respectively. Notice that, the quantity X corresponds to the modulus of $\mathbf{X}$. Moreover, due to symmetry reasons, $\langle F\rangle = |\langle\mathbf{F}\rangle|$ holds for the ensemble average of F, so that measuring the average projection of the force corresponds to measuring the modulus of the average force. Because, in the case of isometric ensemble the force vector is colinear to the end-to-end vector so the time averaged modulus of the force vector is the same as the time averaged of its scalar value. While, in the case of isotensional ensemble, direction of applied force will be alligned to end-to-end vector, due to spherical symmetry, satisfiying the given equality.

One of the possible means to investigate the ensemble equivalence in SMEs is then to analyze the force-extension curves in the two ensembles, namely, of the graphs of $\langle F\rangle\,(X)$ in the isometric ensemble and of $F\,(\langle X\rangle)$ in the isotensional ensemble. In case of ensemble equivalence, these two graphs should, in the limit of infinitely long chains, become indistinguishable. Note, however, that some care has to be taken, about the precise meaning of this statement, as will be discussed in Section 3.1.3.

If the probability density distribution for the end-to-end vector at equilibrium in the free case is denoted by $\mathscr{P}(\mathbf{X})$, then the partition functions for the isometric and isotensional cases can be written, respectively, as

$$Z_{\mathbf{X}} = \mathscr{P}(\mathbf{X}) \tag{3.5}$$

$$Z_{\mathbf{F}} = \int d\mathbf{X}\mathscr{P}(\mathbf{X})\exp(-\beta\mathbf{F}\cdot\mathbf{X}), \tag{3.6}$$

where the formal analogy with the canonical and its conjugate isothermal-isobaric ensemble is evident. Often, as is the case with computer simulations and different theoretical approaches, one needs to refer to a specific molecular model, which will explicitly determine the form of the probability density $\mathscr{P}(\mathbf{X})$. In case of long linear chains, however, the probability density of a

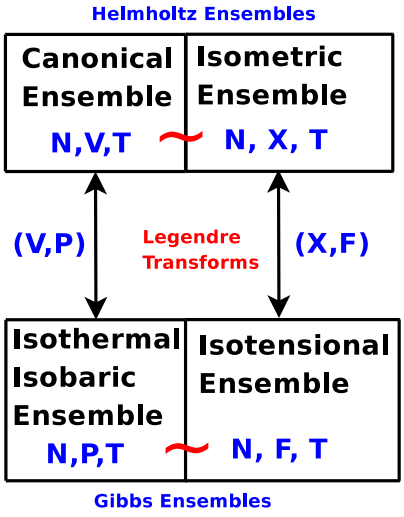


Figure 3.1: The sketch of analogies between conjugate chain ensembles and standard statistical mechanical ensembles.

wide class of models is very well approximated by a Gaussian distribution [145]

$$\mathscr{P}(\mathbf{X}) = b^3 \pi^{-\frac{3}{2}} \exp(-b^2 X^2), \tag{3.7}$$

where the length b^{-1}, is proportional to the root-mean-square end-to-end distance $\sqrt{\langle |\mathbf{X}|^2 \rangle_0} = \sqrt{3}/b$ and to the mean end-to-end distance $\langle X \rangle_0 = 2/(b\sqrt{\pi})$ (the notation $\langle \ldots \rangle_0$ denotes an ensemble average in the free case, i.e. with no applied force). This distribution can be recovered as a limiting case for a wide class of different models representing polymers with discrete units. This is the case, for example, for a Freely Jointed Chain (FJC) of N elements of length a, whose end-to-end distribution is well approximated by the Gaussian distribution when $Na \gg X$. In this case, $b^2 = 3/(Na^2)$ and the total contour length L is $L = Na$. As Neumann noted in a critical analysis on the interpretation of stretching experiments [30], from the partition functions of the Gaussian chain it is possible to derive Hookean-like relations between force and end-to-end vectors. This is easily seen by computing the derivative of the free energy with respect to the end-to-end vector in the isometric case

$$\langle \mathbf{F} \rangle = \beta^{-1} \frac{\partial}{\partial \mathbf{X}} \ln Z_{\mathbf{X}} = 2b^2 \mathbf{X}/\beta \tag{3.8}$$

and computing the average end-to-end vector $\langle \mathbf{X} \rangle$ in the isotensional case

$$\mathbf{F} = 2b^2 \langle \mathbf{X} \rangle / \beta. \tag{3.9}$$

This formal analogy has usually led to the erroneous interpretation of an equivalence between the isometric and isotensional ensembles, which moreover seems to hold for every chain length, and not just in the thermodynamic limit of infinite chain length. In fact, while the scalar version of Equation (3.8) yields $\langle F \rangle = 2b^2 X/\beta$, the isotensional case reads $F = 2b^2 \langle X_{\parallel} \rangle / \beta$, where $X_{\parallel}$ is the projection of the end-to-end vector along the direction of the external force.

Info Box 3 A Random Walk Chain

Consider a one dimensional random walk of N steps, each of length of a_x. Let's assign the number of forward and backward steps that have a Gaussian probability distribution, N_f and N_b respectively. Then end-to-end distance or the net distance travelled forward will be

$$R_x = (N_f - N_b)a_x,$$

and the number of forward states choose out of N steps would be

$$\Omega_x = N!(N_f! - N_b!)^{-1}.$$

If the number of steps is very large, we can use Stirling's approximation

$$lnx! = xlnx - x,$$

so that

$$ln\Omega_x = NlnN - N_f lnN_b - (N - N_f)ln(N - N_f).$$

With a simple change of variable, $f = \frac{N_f}{N}$, on can simplify $ln\Omega_x$ to

$$ln\Omega_x = N\left[flnf + (1-f)ln(1-f)\right].$$

The right hand side of the equation has a minimum at $f = \frac{1}{2}$. If we Taylor expand $ln\Omega_x$ around $\frac{1}{2}$,

$$ln\Omega_x \approx \frac{1}{2}ln\Omega_x + (\frac{1}{2} - f)\left(\frac{d\Omega_x}{df}\right)_{f=1/2} + \frac{1}{2}(1-f)^2\left(\frac{d^2\Omega_x}{df^2}\right)_{f=1/2},$$

then we obtain

$$ln\Omega_x = Nln2 - 2N(\frac{1}{2} - f)^2.$$

Using the fact that $R_x/a_x = N_f - N_b$ and $N = N_f + N_b$ we further can write

$$\Omega_x \sim exp\left(-\frac{R_x^2}{2Na_x^2}\right),$$

It is straight forward to generalize these arguments into 3 spatial dimensions (x, y, z), where $R^2 = R_x^2 + R_y^2 + R_z^2$ and $a_x = a_y = a_z = a$, yielding

$$\Omega_x \sim exp\left(-\frac{3R^2}{2Na^2}\right).$$

Defining the normalization constant c_N as

$$c_N \int_{-\infty}^{\infty} exp\left(-\frac{3R^2}{2Na^2}\right) dR = 1$$

, so that the normalized probability distribution of R will read as

$$P(R) = (2\pi Na^2/3)^{-3/2} exp\left(-\frac{3R^2}{2Na^2}\right).$$

This is valid when $R >> Na$, and can be further simplified with $b^2 = \frac{3}{2Na^2}$ to obtain the final result

$$P(R) = b^3\pi^{-3/2} exp\left(-b^2R^2\right).$$

Info Box 4 ***Isometric Ensemble: Analytical Solution for a rigid model***
[135, 27]
Let us consider rigid model as depicted in Figure 3.3 with a fixed end-to-end vector and fixed bond length. The Hamiltonian of the system of three monomers is

$$H(\theta, p_\theta) = \frac{1}{2}\frac{p_\theta^2}{mr^2},$$

and since the end-to-end distance $2l$ is constant, we can write the partition function of the system at a given temperature T:

$$Z(l,T) = \int\int exp(-\beta H(\theta, p_\theta))d\theta dp_\theta = \int_0^{2\pi}\int_{-\infty}^{\infty} exp(-\beta\frac{1}{2}\frac{p_\theta^2}{mr^2}).d\theta dp_\theta,$$

Recall that $\beta = k_B T$, $\int_{-\infty}^{\infty} exp(-ax^2)dx = \sqrt{\frac{\pi}{a}}$ and $r = \sqrt{a^2 - l^2}$ we get

$$Z(l,T) = 2\pi(2\pi mkT)^{1/2}\sqrt{a^2 - l^2},$$

and analogously the free energy $A(l,T) = -kTlnZ(l,T)$. The force f on the fixed monomers is defined as follows

$$f = \frac{\partial}{\partial(2l)}A(l,T) = \frac{\partial}{\partial(2l)}\left(-kTln(2\pi(2\pi mkT)^{1/2}\sqrt{a^2 - l^2})\right).$$

This expression can be evaluated to

$$f = -\frac{1}{2}kTl(a^2 - l^2)^{-1} = \frac{1}{2}l(l^2 - a^2)^{-1}kT.$$

This relation is confirmed with our Langevin Dynamics simulations given in 3.4.

However, it is clear that these relations cannot hold for the end-to-end distance X. This can already be inferred from the well known fact that in the free case the average end-to-end distance $\langle X\rangle_0$ does not vanish, whereas Equation (3.8) yields zero force at zero distance. A direct evaluation[28] of the $\langle X\rangle$ in the isotensional ensemble, in fact leads to

$$\langle X\rangle = \frac{\langle X\rangle_0}{2}\left[e^{-\nu^2/2} + (\nu + 2/\nu)\int_0^{\nu/2} e^{-t^2}\mathrm{d}t\right], \tag{3.10}$$

where $\nu = \beta F/b$. This relation can be usefully approximated, for small external forces, as $\langle X\rangle = \langle X\rangle_0\left[1 + \frac{1}{12}\nu^2 + \mathcal{O}(\nu^4)\right]$, demonstrating a nonzero extension in the weak force limit in fact for all chain length. Therefore, the force-extension curves measured in the two different ensembles can not coincide. The question of ensemble equivalence has then to be investigated more carefully, defining an observable that quantifies the difference between the two ensembles, and studying its scaling properties as a function of N, i.e., the differences in the finite size effects.

In recent investigations the misinterpretation of Equations (3.8) and (3.9) was certainly overcome, and the general agreement is that in the thermodynamic limit, equivalence between the

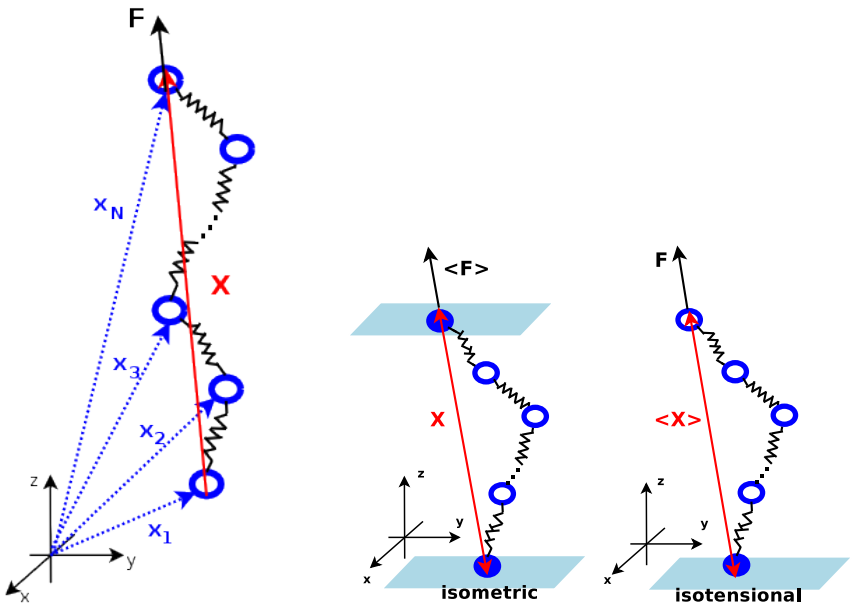


Figure 3.2: Sketch of the chains in the isometric and isotensional ensembles.

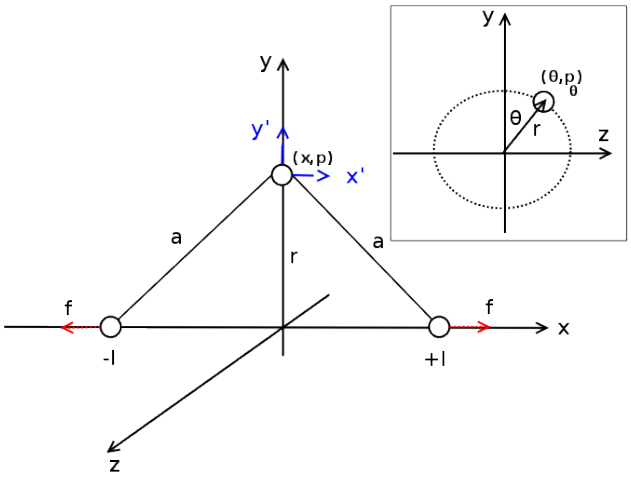


Figure 3.3: Chain of three atoms with both ends fixed, yielding a fixed end-to-end vector (Helmholtz Ensemble). In the rigid model all bond lengths have a fixed value a. In case of the flexible model, the middle atom has harmonic bond with a spring constant of κ. Adopted from [135, 27] .

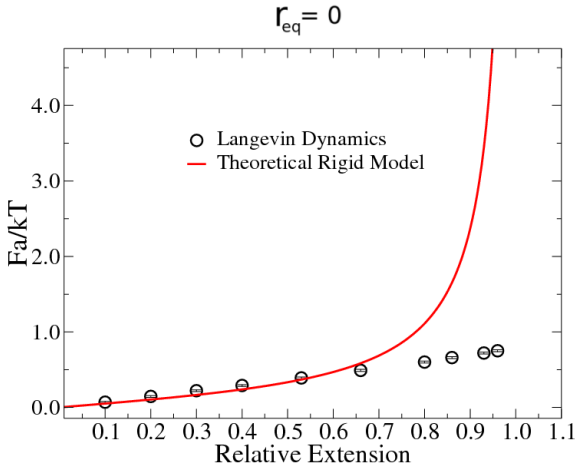


Figure 3.4: Force extension curve for a rigid model chain of three atoms with both ends fixed, fixed end-to-end vector (Helmholtz Ensemble). In the rigid model all bond length have a fixed value, $a = 1$. Langevin simulations were performed with $\kappa = 1.0$ and $r_{eq} = 1$, and $T = 1.0$ and friction constant 1.0. It is seen that in the enthalpic region the theoretical prediction does not hold.

two ensemble is obtained. In particular, employing techniques like renormalization group theory [140], maximum entropy approach [130], and standard analytical techniques [17, 19], all authors conceived the idea that the isotensional and the isometric ensemble are actually equivalent in the usual thermodynamic limit. Recently, however, Neumann pointed out [30] a particular feature of the statistical mechanics of the single chain, namely, that ensemble equivalence cannot be obtained for small values of the external force. In other words, there is theoretical evidence that by choosing an appropriately small external force, the ensemble difference for any (finite, but arbitrarily large) chain length can be maintained constant. To our knowledge, this subtle point in the investigation of the equivalence has never been tested by means of computer simulation, and a part of this chapter is devoted to clarify it, in particular by analyzing the scaling behavior of the ensemble difference as a function of the applied force regime.

3.1.3 Ensemble Difference: Definitions and Scaling

In order to make any statements about equivalence in the thermodynamic limit, there is the need for an observable which quantifies how much two ensembles differ. Given a point (F^*, X^*) on the graph of the isotensional force-extension curve $F(\langle X \rangle)$, the measure of the difference between two ensembles is then defined as

$$\Delta = \frac{X^* - X_{mp}}{X^*}, \tag{3.11}$$

where X_{mp} is the value of the extension in the isometric ensemble that solves the equation $\langle F \rangle (X^*) = F^*$. (The notation stems from the fact that this extension actually coincides with the maximum in the probability distribution function of the isotensional ensemble [17]) A schematic view of the identification of X^* and X_{mp} from the graphs of the isotensional and isometric force-extension curves is given in Figure 3.5. It has to be noted here, that the way the distance between the two ensembles is defined cannot be arbitrary, and in particular it should not depend

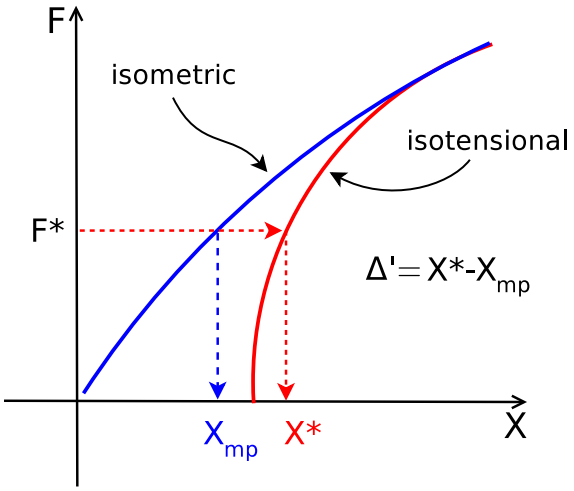


Figure 3.5: The definition of ensemble difference Δ' (non-scaled Δ) between isotensional and isometric ensembles for a given value F^* of the force applied in the isotensional ensemble.

on a reparametrization of the control parameter. So, while $\Delta'(X) = X^* - X_{mp}$ could in principle seem to be a reasonable alternative, it is easily seen that using the same functional form $\Delta'(\xi) = \xi^* - \xi_{mp}$ for the relative extension $\xi = X/N$ would lead to completely different results, while this is not the case when using the relative measure in Equation (3.11).

As it was pointed out by Neumann [30], the correct definition of the ensemble difference shows (at least for the Gaussian chain) that in the moderate or strong stretching regime the ensemble difference indeed goes to zero when the number of monomers tends to infinity, i.e. that ensemble equivalence is obtained. In particular, if the chain end-to-end distance is supposed to scale linearly with the number of monomers in the over-stretched regime, the ensemble difference should scale [17] like $\Delta \sim (N-1)^{-1}$. However, the behavior of the ensemble difference at low stretching regimes is markedly different, since as in this regime both X^* and X_{mp} scale as $\sqrt{N-1}$, and therefore the ensemble difference does not scale with system size, and remains constant. For any given chain length (i.e. for a macroscopic, though not strictly infinite system) it is then possible to find a small enough force for which the ensemble difference does not vanish and, moreover does not decrease appreciably when increasing the chain length. The ensemble difference in the free case Δ_0, as a limiting case of zero forces, can be written as

$$\Delta_0 = 1 - \frac{X_0^*}{X_{mp,0}}.$$

Since for Gaussian chains $X_0^* = 0$, the ensemble difference takes the limiting value of $\Delta_0 = 1$, which for long enough chains is a model-independent result, as long as the Gaussian approximation is valid.

3.2 Langevin Dynamics for Chain Ensembles

The investigations on the ensemble difference have been carried out by simulating the Langevin dynamics of two different chain models by sampling the isometric and in the isotensional ensem-

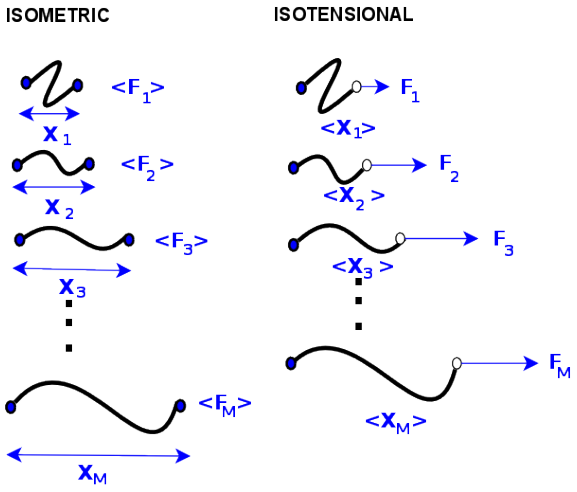


Figure 3.6: For each polymer length, the simulations generate two sets of data, containing M thermodynamic states on the force-extension plane $\{(X_1, \langle F_1\rangle), (X_2, \langle F_2\rangle), (X_3, \langle F_3\rangle), ..., (X_M, \langle F_M\rangle)\}$ and $\{(\langle X_1\rangle, F_1), (\langle X_2\rangle, F_2), (\langle X_3\rangle, F_3), ..., (\langle X_M\rangle, F_M)\}$, for the isometric and the isotensional ensembles, respectively.

bles for different chain length, using the ESPResSo simulation package [61]. Each chain consists of a given number of monomers, which are interacting only with their respective first neighbors via a harmonic potential. Therefore no excluded volume interactions are present and in the free case the chains perform a pure random walk. The role of the Langevin equation

$$\frac{d^2\mathbf{x}_i}{dt^2} = \mathbf{F}_i - \gamma\frac{d\mathbf{x}_i}{dt} + \mathbf{W}_i(t) \tag{3.12}$$

is basically that of providing a thermostat for the chain, where the position of the monomer i is $\mathbf{x}_i$, $\mathbf{F}_i$ represent the conservative force acting on the monomer. As usual, the thermostat is acting via a friction coefficient γ, and a random force $\mathbf{W}_i(t)$ with zero mean and square deviation $\langle \mathbf{W}_i(t) \cdot \mathbf{W}_j(t')\rangle = 6k_BT\gamma\delta_{ij}\delta(t-t')$, in order to satisfy the fluctuation-dissipation theorem [146].

By measuring energies in units of k_BT, distances in arbitrary units d, and considering unitary masses it is possible to express the energy of a chain consisting of N monomeric units, depicted in Figure 3.2, as

$$U = \frac{1}{2}k\sum_{i=1}^{N}(|\mathbf{x}_i - \mathbf{x}_{i-1}| - r_{eq})^2, \tag{3.13}$$

where $k = 20/d^2$ is the spring constant and r_{eq} is the equilibrium distance between a pair of connected monomers.

We focused our attention on two specific cases, namely that of $r_{eq} = d$ and that of zero equilibrium distance. In particular the latter case models precisely a Gaussian probability distribution for the end-to-end distance in the free case. For every simulation, the friction coefficient was set to $\gamma = 0.5/\tau$ and the integration time step was set to 0.01 τ, where $\tau = d(k_BT)^{-1/2}$ is the characteristic time with $k_BT = 1.0$.

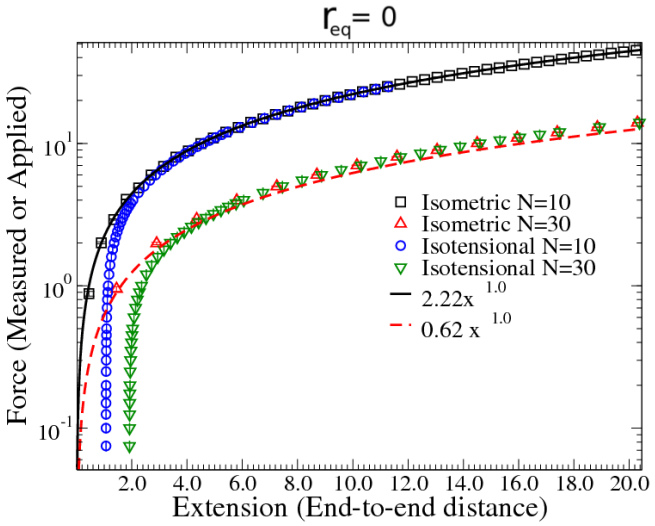


Figure 3.7: The force-extension curve for the $r_{eq} = 0$ case in the isometric (circles) and isotensional (triangles) ensembles. From left to right, $N = 10$, $N = 30$. The dashed lines show the result of linear interpolation.

The sampling of the isometric ensemble was realized by fixing the spatial position of both terminal monomers, x_1 and x_N. In the isotensional ensemble, the end monomer of the chain was fixed, while the other monomers were free to move, and a given force, constant in modulus and direction, was applied to the last one.

The simulation procedures employed for the isometric ensemble consisted of constraining the positions of the first and last monomers at the desired distance X, then performing an relaxation run, up to at least 10 times the auto-correlation time of the observable of interest, F and eventually computing the average $\langle F \rangle = \langle \mathbf{F} \cdot \mathbf{X}/X \rangle$, where $\mathbf{X} = \mathbf{x}_N - \mathbf{x}_1$. In the isotensional ensemble case, the chains started from a straight conformation where the position of the first monomer was constrained, and a constant force $\mathbf{F}$ was applied to the last monomer. Then, after relaxing the system, the end-to-end distance X was sampled. By varying the end-to-end distance over suitable ranges in the isometric ensemble and the magnitude of the applied force in the isotensional ensemble, we sampled the force-extension curves $\langle F \rangle(X)$ and $F(\langle X \rangle)$ for a number of different chain length ranging from 6 to 500 monomeric units. Every point in the force-extension curves was generated form the average taken during a 10^8 steps long run. The procedure is depicted in Figure 3.6.

3.2.1 Force-Extension Curves

In Figures 3.7 and 3.8 we present the measured force-extension curves for the $r_{eq} = 0$ and $r_{eq} = d$ cases, respectively. From the qualitative point of view, every force-extension curve displays the same pattern, namely a difference $\Delta'(X)$ which decreases with increasing applied (or measured) force. This behavior is somewhat expected, and can be qualitatively explained in the following way. In the isotensional ensemble, when the external force is vanishing, $\langle X \rangle$ is expected to have a non-zero average value $\langle X \rangle_0$, the exact value of which depends on r_{eq}. In the isometric case,

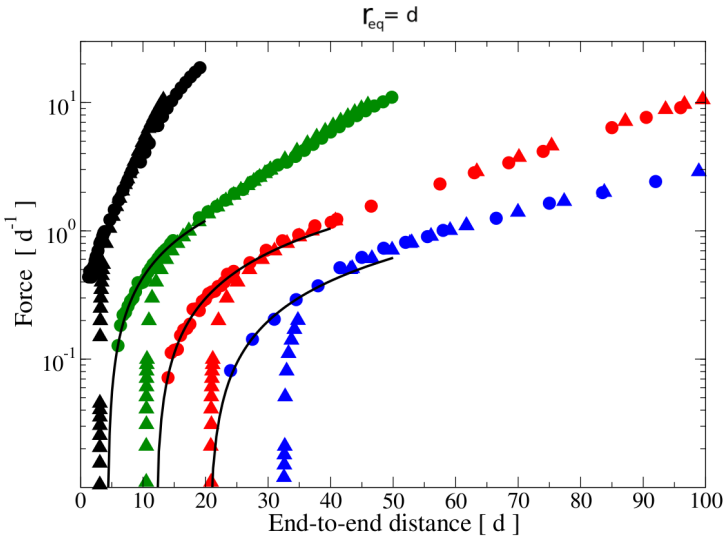


Figure 3.8: The force-extension curve for the $r_{eq} = d$ case in the isometric (circles) and isotensional (triangles) ensembles. From left to right, $N = 10$ (black), $N = 30$ (green), $N = 60$ (red), $N = 90$ (blue).The solid lines represent the result of linear interpolation in the low-stretching regime. Data has been offset along the x-axis for the sake of clarity.

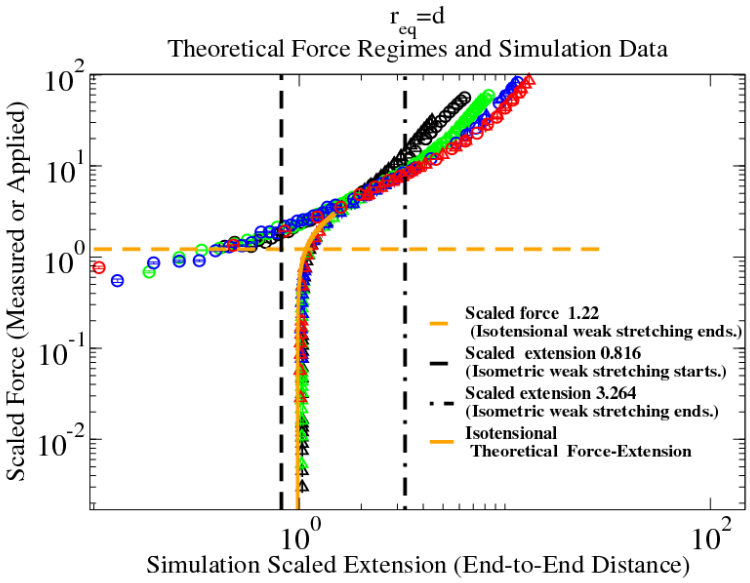


Figure 3.9: Universality of the force-extension curves in the low-stretching regimes for the $r_{eq} = d$ case, reporting the rescaled force F^* as a function of the rescaled end-to-end distance X^* for the isometric (circles) and isotensional (triangles) ensembles. The solid line on top of the isotensional sampled curve in the low-stretching regime is a fit to Equation (3.10), while the horizontal dashed line represents an estimate for the low-stretching regime from dimensional analysis.

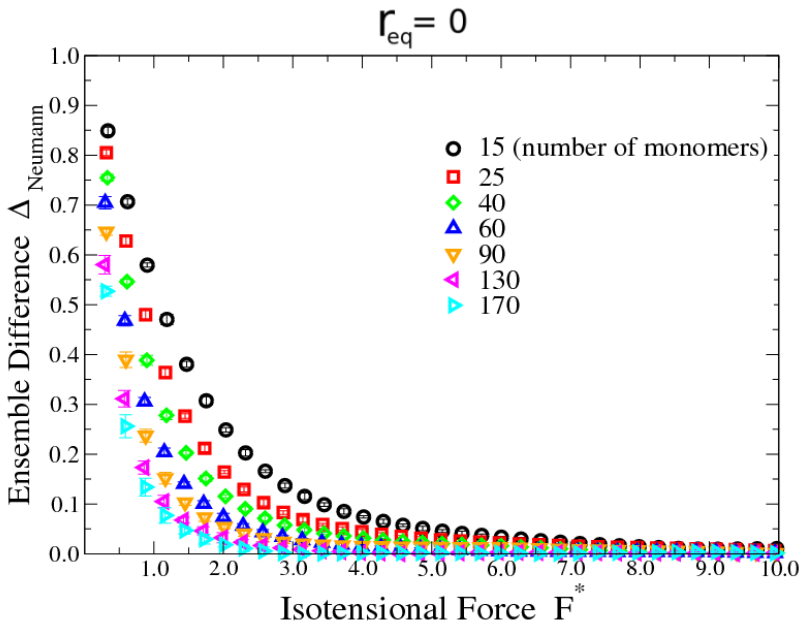


Figure 3.10: Dependence of the ensemble difference Δ as a function of the force applied in the isotensional ensemble, for different chain length.

however, although the projected force $\langle F \rangle$ is not strictly defined at zero end-to-end distance, simple symmetry arguments show that the force acting on each of the terminal beads has, on average, to be zero. Therefore, a vanishing value of $\langle F \rangle$ is expected in this limit. While this arguments account for the differences in the weak stretching regime, in the high stretching regime the energetic contributions to the free energy are expected to dominate over the entropic ones, and therefore a narrowing of the distance between the force-extension curves is expected. In Figure 3.8 the result of a linear fit on the whole spanned x-range is also included, showing the perfect Hookean behavior of chain in the isometric ensemble.

In contrast to the $r_{eq} = 0$ case, the response for the model with a non-zero equilibrium distance does not display a linear behavior over the whole end-to-end distance range. This feature appears because in this case two characteristic distances enter the description of the model, namely, the root mean square displacement of a bead around the equilibrium position, and the equilibrium distance itself. Indeed, since the spring constant k is relatively high with respect to the thermal energy, at low applied forces (or, equivalently, at short end-to-end distances) the chain behaves much like a Freely Jointed Chain (FJC), while when the applied force increases and the springs are significantly stretched with respect to r_{eq}, a different, effective bond length becomes relevant.

Info Box 5 ***The Weak Isotensional Force Law***

Recall that for the weak isotensional force law description the extension can be approximated [28] as follows

$$\begin{aligned} X^* &= X_0^*(1+\frac{1}{12}\nu^2+O(\nu^4)), \qquad (3.14)\\ X_0^* &= 2(b\pi^{1/2})^{-1},\\ b^2 &= 3(2Na^2)^{-1},\\ \nu &= \beta f b^{-1}, \end{aligned}$$

In our simulations we have used $a = \beta = 1.0$ so that the expressions can be simplified as follows

$$\begin{aligned} b^2 &= 3(2N)^{-1} \Rightarrow b = (\frac{3}{2})^{1/2}N^{-1/2} = 1.225N^{-1/2}\\ X_0^* &= 2(1.225N^{-1/2}\pi^{1/2})^{-1} = 2(2.171N^{-1/2})^{-1} = 0.921N^{1/2}\\ \nu &= Fb^{-1} = f(1.225N^{-1/2})^{-1} = 0.816N^{1/2}F. \end{aligned}$$

Now we can define our scaled force and extension as it is presented in Figure 3.9,

$$\widetilde{F} = F(N-1)^{1/2}, \qquad (3.15)$$
$$\widetilde{X} = X(N-1)^{-1/2}, \qquad (3.16)$$

so that the force law in an isotensional ensemble [28, 30], can be written as

$$\begin{aligned} \widetilde{X} &= 0.5X_0^*(\exp(-0.25\nu^2)+(\nu+2\nu^{-1})\int_0^{\nu/2}\exp(-t^2)dt \qquad (3.17)\\ &= 0.46(\exp(-0.16(\widetilde{F})^2)+\\ &\quad (0.81\widetilde{F}+2.44(\widetilde{F})^{-1})erf(0.408\widetilde{F})). \end{aligned}$$

In this way we obtained the theoretical curve given in Figure 3.9, which agrees well with the simulation data in the weak force regime.

The definition of a weak and a high stretching region is of course somewhat arbitrary, although by dimensional analysis, the obvious threshold is set by $F^c = k_BT/a$, where a now identifies the model-dependent effective bond-length. This can be defined [145] according to the scaling behavior of a linear chain in the free case, $a^2 \equiv \langle \mathbf{X}^2\rangle_0/(N-1)$. Since in the weak stretching regime the force is expected to be a small perturbation with respect to the thermal fluctuations, the question naturally arises, whether and to which extent do the scaling arguments hold in the isometric and isotensional ensemble that are valid for the free case only. In order to check this, the force-extension curves have to be rescaled. The end-to-end distance behavior is obviously $X \sim \sqrt{N-1}$, while from Equation (3.9), and remembering that in a Gaussian chain $b \sim 1/\sqrt{N-1}$, one can obtain the scaling behavior for the force $F \sim 1/\sqrt{N-1}$. By plotting the rescaled end-to-end distance and forces, namely,

$$\widetilde{X} = X/\sqrt{N-1}$$

$$\widetilde{F} \;=\; F\sqrt{N-1}$$

for every different chain length, one can see Figure 3.9 that in the low stretching regime both the isotensional and isometric force-extension curves collapse onto the same universal curve. The force-extension curves in the two ensembles fall onto two different universal curves only up to a certain point, after which they start to depart from the universal behavior of free chains and, at the same time, the difference between the isometric and isotensional cases (for a given chain length) diminishes dramatically. In fact, as long as the scaling laws of free chains are fulfilled, they retain the characteristic differences in the force-extension curves. On the other hand, when the chains start being moderately or strongly stretched, the scaling law changes, and the ensemble difference becomes less pronounced.

Info Box 6 ***Force Regimes***

The weak stretching regime is well understood in isotensional and isometric ensembles [28, 29]. For a constant force the weak regime is defined as follows

$$\begin{aligned} 0 &< \nu < 1.0, \\ 0 &< 0.816(N-1)^{1/2}F < 1.0, \\ 0 &< \widetilde{F} < 1.22. \end{aligned}$$

This is the isotensional weak region plotted in Figure 3.9. The most probable end-to-end vector in the zero force limit is

$$X^0_{mp} = 1/b = 0.816(N-1)^{1/2},$$

and the weak stretching regime in the projected isometric ensemble for a given end-to-end distance X [30] with the $(N-1)^{-1/2}$ factor is given by

$$\begin{aligned} X^0_{mp} &< X < 4X^0_{mp}, \\ 0.816 &< \widetilde{X} < 3.264. \end{aligned}$$

This is the isometric weak stretching region plotted in Figure 3.9. The moderate stretching regime is defined above 3.264 *up to a fully stretched conformation.*

3.2.2 Ensemble Inequivalence: Scaling and Conditions

The universal behavior of the chains in the low-stretching regime already points out that the ensemble difference Δ should become constant in this regime, since both $\Delta'(X)$ and X^* scale like $\sqrt{N-1}$. The only requirement for the loss of ensemble equivalence is that the scaled force $\widetilde{F}$ is roughly less than 1.0. This implies that, for example, in order to keep a constant ensemble difference, the applied forces should decrease as $1/\sqrt{N-1}$ with increasing chain length.

Although the scaling arguments are helpful in showing the ensemble inequivalence in the low-stretching regime, it is instructive to explicitly look at the dependence of the ensemble difference on the force, for different chain length and also to look for the scaling behavior of Δ in different force regimes. In Figure 3.10, the ensemble difference is shown as a function of the applied force,

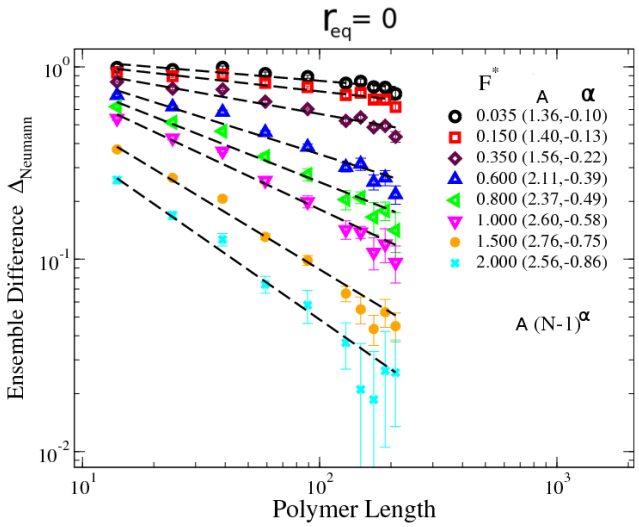


Figure 3.11: Logarithmic plot of the ensemble difference $\Delta_{Neumann}$ as a function of the number of beads, $N-1$, for the $r_{eq}=0$ case. Different symbols correspond to different applied forces in the isotensional ensemble, going from the low-stretching regime (smaller scaling exponent) to the high-stretching regime (larger scaling exponent). Dashed lines correspond to the best fit to a power law. The applied forces ranged from 0.035 to 3.25

.

for different chain length for the $r_{eq}=0$ case. As it has been already noticed from the qualitative analysis of the force-extension curves, the stronger the applied force, the smaller is the ensemble difference. The decay to zero of the ensemble difference happens for smaller forces, the longer the chain is, thus showing that in the moderate stretching regime weaker forces are required to reach the ensemble equivalence. On the other hand, the behavior of the ensemble difference in the weak stretching regime suggests that all the curves tend to converge to the limiting value of 1, no matter what the system size is.

If one focuses instead of the force dependence on the dependence of the number of monomers on the ensemble difference, it is possible to investigate the scaling properties of the ensemble difference in different force regimes. In this way we perform a finite-size study of the convergence of the isometric and isotensional ensembles. In Figures 3.11 and 3.12 the scaling of Δ with respect to the number of units in the chains is shown for the $r_{eq}=0$ and $r_{eq}=d$ cases, respectively.

For both models we observe the same trend, and for each value of the force the ensemble difference actually follows a power-law $\Delta(N) = A(N-1)^{-\alpha}$ (the result of a best fit being also shown in the plots), where the actual value of the exponent strongly depends on the value of the force itself. For diminishing values of the applied force, the curves in the logarithmic plot show a decreasing slope, displaying the clear tendency to become a constant (namely, one) in the limit of vanishing forces. This result not only confirms the inequivalence of the isometric and isotensional ensembles in the vanishing force limit, but also shows the development of the scaling behavior in the intermediate force regime. The limiting behavior in the high-stretching regime can be derived by evaluating the ensemble difference using Equations (3.10) and (3.8), and looking for the asymptotic behavior at large values of ν, leading to $\Delta \sim (N-1)^{-1}$ in the limit of large ν. The two limiting behaviors of the ensemble difference are obviously characterized by the ratio of

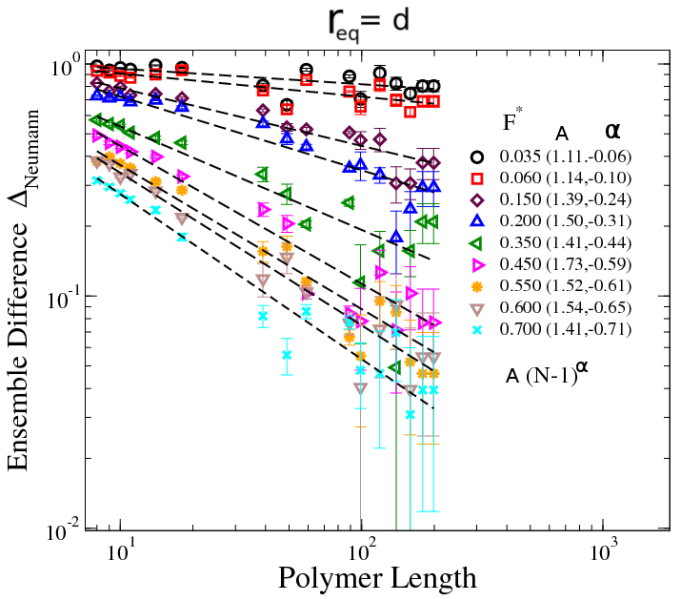


Figure 3.12: Logarithmic plot of the ensemble difference Δ as a function of the number of beads, $N-1$, for the $r_{eq} = d$ case. Different symbols correspond to different applied forces in the isotensional ensemble, going from the lie-stretching regime (smaller slope) to the high-stretching regime (larger slope). Dashed lines correspond to the best fit to a power law. The applied forces ranged from 0.035 to 0.7 d^{-1}

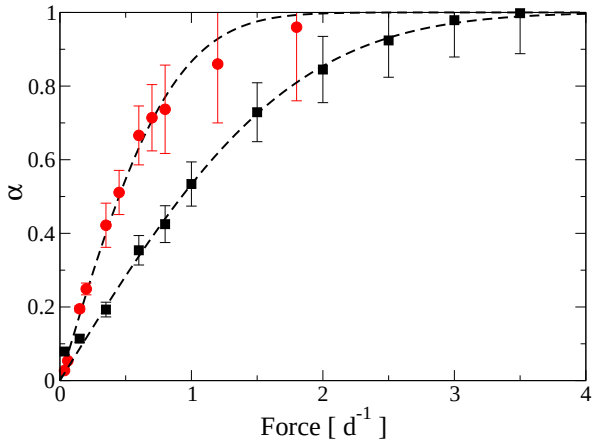


Figure 3.13: Dependence of the fitted scaling exponent α on the applied force for the $r_{eq} = 0$ (squares) and $r_{eq} = d$ (circle) cases. The dashed lines are the result of a best fit to the error function.

the typical energies with the thermal one, $\beta F/b$ which is vanishing and greater then one in the free and overstretched limits, respectively. The way the overstretched region is approached, in terms of scaling exponents will therefore depend on the typical length $1/b$ associated with the model. In Figure 3.13 the value of the scaling exponent of Δ as a function of the applied force is shown for the two models. Indeed, the high-stretching exponent $\alpha = 1$ is reached much faster in the $r_{eq} = 1$ case. A phenomenological fitting function is also reported, which describes surprisingly well the switch between the two regimes is $\alpha(F) = \mathrm{erf}(cF)$, where $c \simeq 0.5$ and 1.0 for $r_{eq} = 0$ and $r_{eq} = d$, respectively (notice that the effective bond length a for the two cases are roughly 0.35 and 1.03).

3.3 Summary

Computer simulations have been employed to study the behavior of linear model polymers in two different ensembles, namely, the isotensional and isometric one, mimicking single molecule experiments. In particular, we have addressed the question of the equivalence of these two ensembles in the thermodynamic limit of growing chain length. Finite size effects of order $(N-1)^{-1}$ for the ensemble difference have been demonstrated as a limiting case for a Gaussian chain experiencing high tensions, and verified for bead-spring models with zero and non-zero equilibrium distance. In this case ensemble equivalence is reached in the infinite chain length limit. By switching to the low-stretching regime, a dramatic change in the scaling behavior appears. Here we find that the force-extension curves exhibit an universal scaling behavior that is typical for free Gaussian chains in equilibrium. This, in turn, leads to the fact that ensemble equivalence can indeed never be obtained in the vanishing force limit, as has been pointed out by Neumann [30]. Our computer simulations confirm and enhance the analysis by Neumann, showing that care has to be taken when considering the thermodynamics of single molecules, which presents many subtle differences with respect to bulk systems, despite the formal analogies between the two systems.

4 Induced Charge Computation

In this chapter we address the problem of how to compute efficiently electrostatic interactions between charged particles in the presence of arbitrarily shaped dielectric domains. In Section 4.1, the general motivation of continuum electrostatics will be summarized. In Section 4.2 the basic form of Poisson equation for inhomogeneous dielectrics will be discussed. A conceptual framework for computing the polarization charge density on the dielectric interface will be given in Section 4.3. The matrix formulation which is called *Induced Charge Computation* (ICC), will be discussed, too. Furthermore, an iterative approach to solve the ICC system of linear equations will be presented in Section 4.4. It will be demonstrated that the problem can be handled without solving the whole boundary value problem formally. Details of our implementation and test cases for different sample geometries will be presented in Section 4.6. The computational comparison of the ICC method and the iterative algorithm ICCP^{3}M will be given in Section 4.7.

4.1 Continuum Electrostatics: A Perspective

Independent of the origin of the problem, be that physics, chemistry, or biology, as well as in computational mathematics and electronics engineering, computing electrostatic forces is an important generic problem [147, 45, 148, 149, 47]. To solve the Poisson equation and its variants in the context of N-body problems turns out to be a computational challenge in molecular simulations in general [150, 151]. A particularly rich set of problems appears in soft matter and biophysics, when spatially different dielectric domains are present.

The dielectric continuum (DC) model from biophysics point of view has been used in the literature as an approximation to solvent degrees of freedom in molecular and related systems [41, 47]. The primary idea of the method is to replace solvent particles with a dielectric background that changes the electrostatic interaction strength among solute particles. The DC approximation is quite advantageous in situations where the explicit solvent dynamic time scale is much smaller then that of the the model system [152]. Otherwise the computational overhead to relax the many order of magnitute more numerous, but rather interesting solvent degrees of freedom would be too dominant, and prohibitive to treat the interesting length and time scales.

In recent times, DC models for solvents have become increasingly popular and are widely used [147, 45, 148]. In these investigations, the electrostatic energy is obtained by solving numerically Poisson equation (the Poisson-Boltzmann equation in the presence of ions, Poisson-Nernst-Planck equation if an ionic current is considered) with various techniques, such as the finite difference method [153, 154, 155, 156], surface boundary elements [55, 59, 157, 158, 159], finite elements [160], multidimensional optimization [50], surface-charge variational approach [48, 57, 161, 162, 163, 58, 164, 165, 166, 167, 158, 168], matrix compression technique [169, 170], Maggs method [171, 172], Green function [47], Generalised Born technique [173], and the volume polarization approaches [174].

It is shown that in realistic simulations of ionic channels and related macromolecular transport through pores, Dielectric Boundary Force (DBF), the force occur due to dielectric boundary

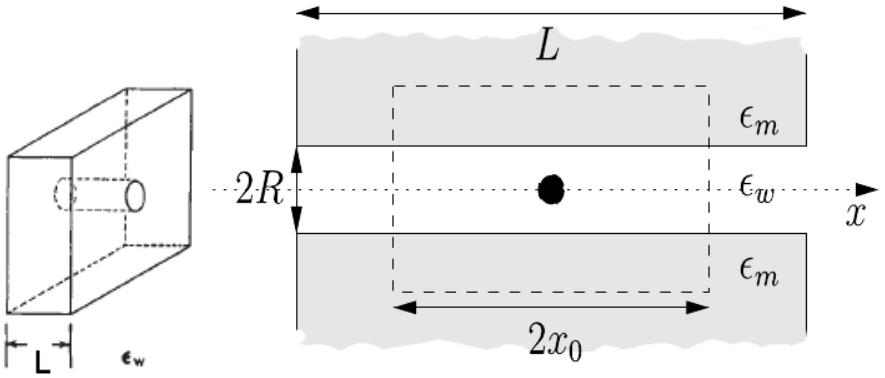


Figure 4.1: A polarizable cell membrane acts as a barrier to the passage of solute material between two aqueous compartments. Electrostatic energy of this model discussed first by Parsegian [41]. Schematic view of a channel [175] of length L, radius R with a charge e (an ion) sitting in its center, i.e. at x = 0 is shown. The channel consists of water with, of dielectric constant $\epsilon_w \approx 80$, and is surrounded by the membrane, of dielectric constant $\epsilon_m \approx 2$. A typical ion channel, e.g. gramicidin A, has a length $L \approx 2.5$ nm and a radius R ≈ 3 Å.

present in the system, plays a dominant part due to the difference in the dielectric constants of the solvent and the membrane [43, 42].

One particular problem of this kind is, for example a polarizable cell membrane, that is acting as a barrier to the passage of ions or charged macromolecules between two aqueous compartments [41, 175] as depicted in Figure 4.1. One apparent obstacle to ionic conduction is the electrostatic energy barrier which has to be passed in order to move the ion through the channel. In 1969, Parsegian [41] found that, for an infinite channel, the electrostatic self-energy E, i.e. the energy required to move an ion from infinity to the center of the channel, is of the order of 16 k_{BT}. A major contribution to this self-energy is due to the consequent difference between the dielectric constant of the water channel ($\epsilon_w \approx 80$, at room temperature) and that of the lipid membrane ($\epsilon_m \approx 2$). Such a large barrier would prevent any ionic flow, contrary to the relatively high conductances observed in experiments [4]. The most obvious limitation of previous models is that they consider channels of infinite length (no screening by counter-ions is assumed). Taking into account the finite length of the channel (gramicidin A has a characteristic length of 25 Å), a dielectric shield of the channel, a considerable decrease of this barrier, to $6.7 k_B T$, is needed [175].

In a finite channel, shown in Figure 4.1, at distances x, along the channel, $x_0 \ll x$, the electrostatic potential is three dimensional. In the region $x_0 \gg x$, the ions are subject to a constant repulsive force and in the one dimensional Coulomb regime. x_0 is the threshold length where electrostatic potential changes its behaviour, 3D to 1D. For large lengths $L \gg 2x_0$, Parsegian has proposed the self-energy for a finite pore system shown in Figure 4.1 :

$$E^L = E^\infty - \frac{e^2}{L\epsilon_m} \log\left(\frac{2\epsilon_w}{\epsilon_w + \epsilon_m}\right),$$

where the first term is the self-energy of the infinite channel and the second term corresponds to the self energy of a charge membrane without a pore. It is evident from the formula that the larger the difference between the dielectric constants, ϵ_w and ϵ_m, is the more pronounced will be the electrostatic energy an ion experience due to its self-energy in the pore.

4.2 Poisson Equation: Dielectric Interfaces

The Poisson equation is an elliptic partial differential equation (PDE). It describes the electrostatic potential $\phi(\mathbf{r})$ of a continuous distribution of electric charges located at $\mathbf{r}$ with a local charge density $\rho(\mathbf{r})$ confined to a domain $\mathscr{D}$ of volume V, with a boundary $\mathscr{B}$. Two common situations are: one with infinite boundary or a periodically repeated one. For a continuous distribution of charges in a free-space, i.e. not conducting or dielectric bodies, Poisson equation in CGS system of units satisfies

$$\nabla^2\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}), \tag{4.1}$$

subject to suitable boundary conditions. Dirichlet boundary condition is chosen here. It means that the potential is specified everywhere except on the boundary [176]. The closed form analytic solution of Poisson Equation (4.1) can be expressed in terms of the Green function, where it is defined as follows for Laplacian operator ∇^2

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|},$$

hence the following potential expression solves the Poisson equation (4.1)

$$\phi(\mathbf{r}) = \int_{\mathscr{D}} d\mathbf{r}'\rho(\mathbf{r})G(\mathbf{r}, \mathbf{r}'). \tag{4.2}$$

If system properties, such as dielectric properties are not homogeneous we have to deal with a Poisson equation, where the strength of the potential will be coupled to a distance dependent dielectric function. This situation could be quantitatively characterized by $\epsilon(\mathbf{r})$, a dielectric constant as a function of distance, and an electric potential $\phi(\mathbf{r})$, that satisfies the modified Poisson equation. Hence, the Poisson equation for vacuum given in Equation (4.1) that is modified for inhomogeneous dielectrics as follows,

$$\nabla(\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})) = -4\pi\rho(\mathbf{r}). \tag{4.3}$$

Applying the product rule of elementary vector calculus on the left hand side, we obtain

$$(\nabla\epsilon(\mathbf{r}))(\nabla\phi(\mathbf{r})) + \epsilon(\mathbf{r})\nabla^2\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \tag{4.4}$$

Dividing both sides of this equation by $\epsilon(\mathbf{r})$ and factoring the remaining expressions with 4π will yield the following expression for $\nabla^2\phi(\mathbf{r})$

$$\nabla^2\phi(\mathbf{r}) = -4\pi\rho_{eff}(\mathbf{r}), \tag{4.5}$$

where the effective charge density $\rho_{eff}(\mathbf{r})$ is the remaining expression from our re-arrangement operation in Equation (4.4), hence

$$\rho_{eff}(\mathbf{r}) = \left[\frac{\rho(\mathbf{r})}{\epsilon(\mathbf{r})} + \frac{\nabla\epsilon(\mathbf{r})}{4\pi\epsilon(\mathbf{r})}\nabla\phi(\mathbf{r})\right] \tag{4.6}$$

This re-arrangement gives us an opportunity to express the electric potential $\phi(\mathbf{r})$ in the Poisson equation in Equation (4.3), using similar Green function expression given in Equation (4.2).

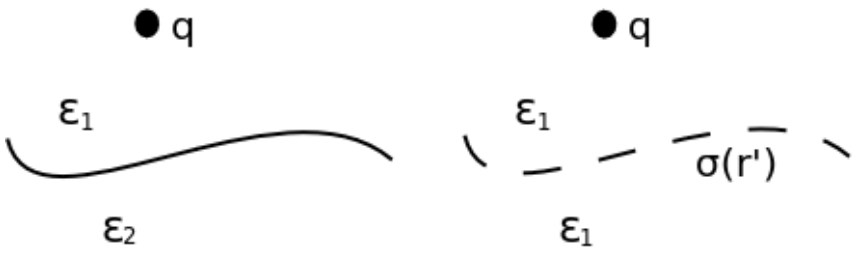


Figure 4.2: The model problem depicted on the left, two dielectric domains ϵ_1 and ϵ_2 are separated by an interface and a source charge q sitting on one region. The concept of polarization charge density, $\sigma(r')$ suggests the following: one can replace the interface separating two dielectric regions by a surface charge density and embed the whole system into one dielectric domain, of ϵ_1, which is typically the one where the source charges reside. This situation is depicted on the right.

Hence, the electric potential $\phi(\mathbf{r})$ in this case will read,

$$\begin{aligned}\phi(\mathbf{r}) &= \int_{\mathscr{D}} d\mathbf{r}' \rho_{eff}(\mathbf{r}') G(\mathbf{r},\mathbf{r}') \\ &= \int_{\mathscr{D}} d\mathbf{r}' \frac{\rho(\mathbf{r}')}{\epsilon(\mathbf{r}')|\mathbf{r}-\mathbf{r}'|} \\ &\quad + \int_{\mathscr{D}} d\mathbf{r}' \frac{\nabla\epsilon(\mathbf{r}')}{4\pi\epsilon(\mathbf{r}')} \frac{\nabla\phi(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \end{aligned} \tag{4.7}$$

This formulation will be the starting point of an electrostatic problem involving dielectric domains, subject to appropriate boundary conditions. Notice that by this re-arrangement we have naturally split the charge density into two components in Equation (4.7) and obtain a Poisson equation with $\rho_{eff}(\mathbf{r})$ in Equation (4.5),

$$\rho_{eff}(\mathbf{r}) = \rho(\mathbf{r})/\epsilon(\mathbf{r}) + \rho_{pol}(\mathbf{r}). \tag{4.8}$$

This approach is not only plausible from mathematics point of view but also physical interpretation, where $\rho(\mathbf{r})$ corresponds to the charges embedded in the media and $\rho_{pol}(\mathbf{r})$ denotes the polarization charges on the dielectric interface. We will discuss polarization charges in the next section.

4.3 Surface Polarization Charge

The concept of polarization charge is a well known result of classical electrostatics. The polarization effects of the free volume charges can be mimicked as an induced surface charge density on the interface that separates two continuum dielectric domains (homogeneous and isotropic), such as a solvent and a membrane [55, 46, 51, 56, 170].

The introduction of a polarization charge distribution on the dielectric boundary that is the result of the difference in polarization in the different dielectric domains, with the entire space now assumed to be filled with a single dielectric. The electrostatic problem involving different dielectric domains and true charges can be treated as a problem of a true and polarization charges in a single dielectric domain [55]. This situation is depicted in Figure 4.2. If the polarization in

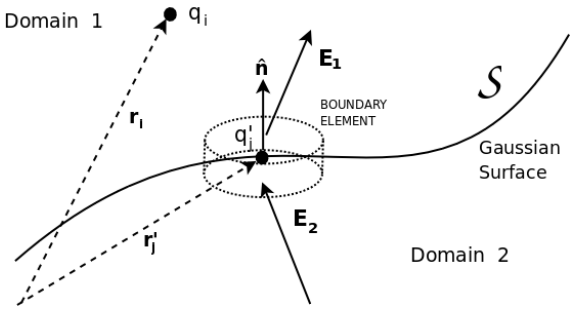


Figure 4.3: A sketch of the dielectric domains ϵ_1 and ϵ_2 separated by a Gaussian surface. The source charges $q_1, q_2, ..., q_N$ are sitting in one domain at $\mathbf{r}_i$, $i = 1, ..., N$. These charges give rise to an induced surface charges $q_j^{'}$, $j = 1, ..., M$ which are located at $\mathbf{r'}_j$, $j = 1, ..., M$ in the center of discretized boundary elements, respectively. Due to the discontinuity of the field at the boundary of two domains, $\mathbf{E}_1$ and $\mathbf{E}_2$ appear. If we don not consider a dielectric surface but use the surface polarization as described in the text, we have a continuous field $\mathbf{E}$ over induced charges and source charges embedded in a single domain of dielectric ϵ_1.

a dielectric domain is uniform, the polarized volume charge does not contribute to the field, but only the surface polarization charge at the boundary of dielectric domain does. Then, the surface charge density is determined as the result of applying boundary conditions.

Consider two dielectric domains of dielectric constants ϵ_1 and ϵ_2 respectively. Let the set of discrete charges q_i, $i = 1, .., N$ be located in the dielectric domain with dielectric constant ϵ_1. The positions of these charges are $\mathbf{r}_i$, $i = 1, .., N$ respectively. This situation is depicted in Figure 4.3. We focus on the case, where two dielectric domains are divided by a smooth surface of an arbitrary shape. The primary idea of *polarization charges*, as we mentioned, is that the effect of different dielectric domains in computing the electric field will be represented by surface charges, $q_j^{'}$, $j = 1, .., M$, which are located at $\mathbf{r'}_j$, $j = 1, .., M$ and where the whole system is now embedded in a single medium of dielectric constant ϵ_1. Now, let us come back to the electric potential given in Equation (4.7). The first term of the right hand side is nothing but the sum of Coulomb potentials due to source charges, let's call this potential $\phi_{source}(\mathbf{r})$, and its discrete counter part reads

$$\phi_{source}(\mathbf{r}) = \frac{1}{\epsilon_1} \sum_{i=1}^{N} \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|}. \tag{4.9}$$

Notice that the electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ are defined in the original problem before the introduction of surface polarization and defined to be:

$$\mathbf{E}_1 = -\nabla\phi(\mathbf{r})|_{Domain1}\,, \tag{4.10}$$
$$\mathbf{E}_2 = -\nabla\phi(\mathbf{r})|_{Domain2}\,. \tag{4.11}$$

The second term on the right side of Equation (4.7) demands for a more accurate explanation. Equation (4.7) describes the potential, with the following boundary conditions:

$$\epsilon_1 \mathbf{E}_1|_{Domain1} \cdot \hat{\mathbf{n}} = \epsilon_2 \mathbf{E}_1|_{Domain2} \cdot \hat{\mathbf{n}}. \tag{4.12}$$

Hence, the relations between $\mathbf{E}$ and the set $\mathbf{E}_1$, $\mathbf{E}_2$:

$$\mathbf{E}_1 \cdot \hat{\mathbf{n}} = \mathbf{E} \cdot \hat{\mathbf{n}} - 2\pi\sigma(\mathbf{r}') \cdot \hat{\mathbf{n}}, \tag{4.13}$$

$$\mathbf{E}_2 \cdot \hat{\mathbf{n}} = \mathbf{E} \cdot \hat{\mathbf{n}} + 2\pi\sigma(\mathbf{r}') \cdot \hat{\mathbf{n}}, \tag{4.14}$$

where $\mathbf{E}$ is the continuous field when induced surface charges and real charges are embedded in a single dielectric continuum. This field implicitly defines the force acting on the boundary element as $\mathbf{F}_j = \mathbf{E}_j/q'_j$, $q' = \sigma(\mathbf{r}'_j)a_j$, where a_j is the jth boundary element area. The expression $2\pi\sigma(\mathbf{r}')$ is the field generated but the induced charge located at $\mathbf{r}'_j$. The polarization surface charge density $\sigma(\mathbf{r}')$ is the observable that we are trying to determine.

Finally, the second part of the total electric potential given in Equation (4.7) can be presented on the boundary as follows, using that $\nabla\epsilon(\mathbf{r}') = \epsilon_1 - \epsilon_2$ and $\epsilon(\mathbf{r}') = \frac{\epsilon_1+\epsilon_2}{2}$:

$$\phi_{polarize}(\mathbf{r}') = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} \int_{Interface} da\hat{\mathbf{n}} \cdot \nabla\phi(\mathbf{r}) \frac{1}{2\pi|\mathbf{r} - \mathbf{r}_i|}. \tag{4.15}$$

Notice, that $\phi_{polarize}(\mathbf{r}')$ is a surface integral, $\nabla\epsilon(\mathbf{r})$ is zero everywhere but on the surface that divides the two dielectric domains.

Info Box 7 ***Solving Poisson Equation with Finite Differences Numerically*** *[52]*
Consider the Poisson equation in 2D for an electric potential $u(x, y)$,

$$u_{xx} + u_{yy} = -1$$

defined on a square region $0 \leq x \leq 1$ and $0 \leq y \leq 1$, subject to condition $u = 0$ on the boundary. One can cover the region with grids of size $h = M^{-1}$ for an integer M. The example is shown in Figure 4.4. The 5-point difference equation reads

$$h^{-2}\left[u(x_i + h, y_i) + u(x_i - h, y_i) + u(x_i, y_i + h) + u(x_i, -y_i) - 4u(x_i, y_i)\right] = -1$$

where the index i runs over all grid points. This equality generates a system of linear equations to be solved with standard linear algebra techniques. But when the system size (number of grids) increases it becomes computationally expensive to solve the large system of equations.

4.4 Boundary Element Method

Here we will show how to derive the iterative scheme in a simple way without using the complex approaches studied in [48, 164]. The procedure is quite similar to the approach previously proposed in [51] based on the fast multipole method.

The continuity of the potential on the surface is guaranteed by boundary conditions on the dielectric interface, given in Equation (4.12). We can reach an expression for the surface charge

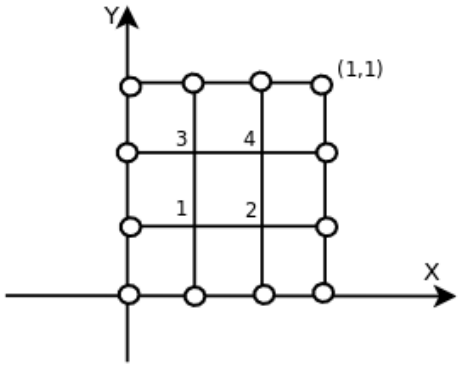


Figure 4.4: A finite difference scheme for a Poisson equation on a domain, that is divided by rectangular grids. Local potential values at points 1,2,3 and 4 are determined via a finite differences scheme. Elements shown with circles are boundary grids, where the values are fixed.

density that is discretized with induced charges on the boundary elements by using Equations (4.13) and (4.13):

$$\epsilon_1(\mathbf{E}\cdot\hat{\mathbf{n}} - 2\pi\sigma(r')) = \epsilon_2(\mathbf{E}\cdot\hat{\mathbf{n}} + 2\pi\sigma(r')), \tag{4.16}$$

$$(\epsilon_1 - \epsilon_2)\mathbf{E}\cdot\hat{\mathbf{n}} = (\epsilon_1 + \epsilon_2)2\pi\sigma(r'), \tag{4.17}$$

$$\sigma(r') = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}\frac{1}{2\pi}\mathbf{E}\cdot\hat{\mathbf{n}} = f\mathbf{E}\cdot\hat{\mathbf{n}}, \tag{4.18}$$

where f is a constant depends on the dielectric constants

$$f = \frac{1}{2\pi}\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}.$$

The normal component of the continuous electric field at the test point with discretized surface charge consist of two terms:

$$\mathbf{E}\cdot\hat{\mathbf{n}} = \mathbf{E}_{induce}\cdot\hat{\mathbf{n}} + \mathbf{E}_{source}\cdot\hat{\mathbf{n}}, \tag{4.19}$$

one describes the contribution of the boundary element $\mathbf{E}_{induce}$, and the contribution of the source charges $\mathbf{E}_{source}$. The similarity of this equation and Equation (4.8) shows that the electric potential can be split into two components if two dielectric media are considered.

4.4.1 Equivalent Matrix Formulation leading to ICC

To characterized the position of induced charges two indexes are used, namely j and k. For source charges we used the index i, recall Figure 4.3. We can write the electric field on the given induced charge j

$$\mathbf{E}_{induce}\cdot\hat{\mathbf{n}}^j = K_{jk}\sigma_j, \tag{4.20}$$

and the electric field on the given induced charge due to all source charges

$$\mathbf{E}_{source}\cdot\hat{\mathbf{n}}^j = \frac{1}{\epsilon_1}\sum_i \frac{1}{|\mathbf{r}_i - \mathbf{r'}_j|}\hat{\mathbf{r}}_{ij}\cdot\hat{\mathbf{n}}^j. \tag{4.21}$$

Then the following calculations are possible for Equation (4.19), by using the fact that $\mathbf{E} \cdot \hat{\mathbf{n}}^j = \frac{1}{f}\sigma^j(\mathbf{r'}_j)$ from Equation (4.18),

$$\begin{aligned}
\frac{1}{f}\sigma_j(\mathbf{r'}_j) &= \mathbf{E}_{induce} \cdot \hat{\mathbf{n}}^j + \sum_i \mathbf{E}_{source} \cdot \hat{\mathbf{n}}^j; && (4.22)\\
\frac{1}{f}\sigma_j(\mathbf{r'}_j) &= K_{jk}\sigma^j(\mathbf{r'}_j) + \sum_i \mathbf{E}_{source} \cdot \hat{\mathbf{n}}^j; && (4.23)\\
\sigma^j(\mathbf{r'}_j) - fK_{jk}\sigma^j(\mathbf{r'}_j) &= f\sum_i \mathbf{E}_{source} \cdot \hat{\mathbf{n}}^j; && (4.24)\\
(\delta_{jk} - fK_{jk})\sigma^j(\mathbf{r'}_j) &= f\sum_i \mathbf{E}_{source} \cdot \hat{\mathbf{n}}^j, && (4.25)
\end{aligned}$$

where δ_{jk} is the Kronecker delta. The explicit form of the elements of matrix K is

$$K_{jk} = \begin{cases} \sum_k \frac{1}{|\mathbf{r}_j - \mathbf{r}_k|}, & \text{if } j \neq k, \\ 0, & \text{if } j = k. \end{cases}$$

Note that Equation (4.25) is nothing but the equivalent version of the matrix formulation derived by Boda et al. [58] via density functional formulation [48],

$$\begin{aligned}
\mathbf{Ah} &= \mathbf{c}, && (4.26)\\
\mathbf{h} &= \mathbf{A}^{-1}\mathbf{c},
\end{aligned}$$

which is called Induced Charge Computation (ICC) algorithm. From our basic derivation we can identify the specific form of the vectors h and c. and the matrix A as:

$$\begin{aligned}
A_{jk}h_j &= c_j, && (4.27)\\
A_{jk} &= (\delta_{jk} - fK_{jk}),\\
h_j &= \sigma^j,\\
c_j &= f\sum_i \mathbf{E}_{source} \cdot \hat{\mathbf{n}}^j.
\end{aligned}$$

Hence, the problem of computing polarization charge density turns out to be a linear problem of the form, given in Equation (4.27).

4.5 ICCP^{3}M : Bharadwaj-Tyagi Algorithm

Solving a sparse system of linear equations in the form $\mathbf{Ax} = \mathbf{b}$, for example such as one formulated as in Equations (4.26) and (4.27) for solving dielectric problem derived with different procedures, is a well founded and probably the most studied problem in numerical analysis [177]. However, techniques such as Gauss elimination [177] is not well suited for large sparse systems of linear equations, due to its large memory storage requirements and its computational complexity. An alternative approach is provided by iterative methods [52].

The central idea of the iterative approach is formulated as an alternative equivalent matrix equation. The unknown vector $\mathbf{x}$ can be expressed in terms of itself, a known coefficient matrix $\mathbf{A}$ and

a known vector $\mathbf{b}$. Hence, one can obtain set of iterative linear equations for the unknown vector $\mathbf{x}$. At this point, the procedure can be executed on a component level without the need of storing all the components of a known coefficient matrix $\mathbf{A}$. After the initial guess of the unknown vector, the iterative procedure updates the unknown vector with new values. The iteration can be stopped when the required accuracy is achieved.

The simplest type of iterative approach is called *Jacobi*. However, one can make this scheme a little smarter by introducing computed components instantly within the same iteration. Furthermore introducing a relaxation parameter to increase the convergence rate have been used as a standard tool for iterative schemes, such as the Successive Over-relaxation (SOR) method [177, 52].

Info Box 8 ***Iterative Approach in Solving Linear Systems***

To solve the Equation (4.25) for large number of charges is a linear problem of the form:

$$\mathbf{A}x = b.$$

One of the most computationally efficient methods to solve it is to use an iterative scheme. The basic idea of this method, see also relaxation methods for boundary value problems Ch.19.5 [177] and the work in [52] for general review, is to present the matrix equation in the following form that one can construct an iterative equation over the unknown variables :

$$\mathbf{B}x^{old} + b = x^{new},$$

where $\mathbf{B}$ *is a matrix* $\mathbf{B} = \mathbb{I} - \mathbf{A}$*. The simplest iteration scheme called* Jacobi *can be written in component notation as*

$$x_j^{n+1} = B_{ji}x_i^n + b_i,$$

where x_j^{n+1} *and* x_i^n *are the new and old values respectively. Similarly we can use the values of* x *as long as it is available this is called* Gauss-Seidel*, beside the fact that it uses new values of* x_i *as soon as available. More efficient method is to introduce a relaxation parameter* ω *(the value of which is proven to vary from zero to two [178]) is called* Successive Over-relaxation method *(SOR) and has the following form:*

$$x_i^{n+1} = \omega(B_{ji}x_i^k + b_i) + (1 - \omega)x_i^n.$$

More complex schemes exist but we only cover SOR here, because we will implement it in the ICCP^3*M algorithm in this thesis.*

A standard Successive Over-relaxation (SOR) scheme of equivalent form given in Equation (4.27), that is appropriate for iterative solution of induced charges σ_j can be written as follows:

$$\sigma_j^{n+1} = \omega\left((\delta_{jk} - A_{jk})\sigma_j^n + c_j\right) + (1-\omega)\sigma_j^n, \tag{4.28}$$

$$\sigma_j^{n+1} = \omega\left((\delta_{jk} - \delta_{jk} + fK_{jk})\sigma_j^n + f\sum_i \mathbf{E}_{source}\cdot\hat{\mathbf{n}}^j\right) + (1-\omega)\sigma_j^n, \tag{4.29}$$

$$\sigma_j^{n+1} = \omega\left(fK_{jk}\sigma_j^n + f\mathbf{E}_{source}\cdot\hat{\mathbf{n}}^j\right) + (1-\omega)\sigma_j^n, \tag{4.30}$$

where ω is the relaxation constant, that accelerates convergence, superscript $n+1$ corresponds to new value and n old value of given induced charge. We can use the fact that

$$f\mathbf{E}\cdot\hat{\mathbf{n}}^j = fK_{jk}\sigma_j^n + f\mathbf{E}_{source}\cdot\hat{\mathbf{n}}^j.$$

Hence Equation (4.30) can be expressed as follows:

$$\sigma_j^{n+1} = \omega(f\mathbf{E}\cdot\hat{\mathbf{n}}^j) + (1-\omega)\sigma_j^n. \tag{4.31}$$

Recall that the field $\mathbf{E}$ in Equation (4.31), is the electric field due to both induced charges and the source charges embedded in a single dielectric media, dielectric constant of ϵ_1. This field can be computed with any standard Coloumb solver, **without any dielectric jump**. The derived matrix Equation (4.25), for solving Poisson Equation with two dielectric medium can be solved via the SOR approach. The expression for dielectric SOR is given in Equation (4.31),

$$\sigma_j^{n+1} = \omega(f\mathbf{E}_{charges}\cdot\hat{\mathbf{n}}) + (1-\omega)\sigma_j^n, \tag{4.32}$$

re-arranging this equation, yields the Induced Charge Computation (ICC)

$$\sigma_j^{n+1} = \sigma_j^n + \omega\left(f\mathbf{E}_{charges}\cdot\hat{\mathbf{n}} - \sigma_j^n\right) \tag{4.33}$$

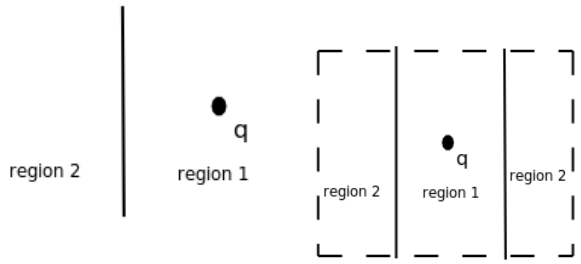
We can get $\mathbf{E}_{charges}$ via the force acting on the charged particle computed with the P^{3}M , which a fast and efficient $N-body$ electrostatic force computation algorithm, that is a well understood in the context of charged systems [60]. So that, we have utilized P^{3}M in computing $\mathbf{F}^{P3M}_{charges}$ on the charges (q_j) and the areas (a_j), so that Equation (4.33) will read

$$\frac{q_j^{n+1}}{a_j} = \frac{q_j^n}{a_j} + \omega(f\mathbf{F}^{P3M}_{charges}\cdot\hat{\mathbf{n}}_j - \frac{q_j^n}{a_j}) \tag{4.34}$$

The procedure can be launch with a random surface charge distribution on the given dielectric boundary and can be repeated until a convergence criterion Δ_{conv} has been reached, such that the maximum charge difference that is generated by the latest iteration step $n+1$ does not deviated more then Δ_{conv} compared to the value from the previous step n, algebraically

$$max\left(abs(q_j^{n+1} - q_j^n)\right) < \Delta_{conv}.$$

The value of Δ_{conv} should not be smaller then the P^{3}M accuracy, otherwise convergence will not be achieved.



(a) Configuration of the problem for ICC.

(b) Configuration of the problem for ICCP^{3}M .

Figure 4.5: A single charge sitting in the region 1 in front of a dielectric interface of planar shape that divides region 1 from a different dielectric domain region 2. For ICCP^{3}M , due to periodic boundary conditions, a second planar dielectric interface must be introduced to avoid an ill definition of the domains on the box boundaries. This is an approximation to the problem.

4.6 Numerical Results and Comparisons

So far we have discussed the formulation of algorithms and their working principles. In this section test examples will be discussed and, when it is possible the comparison to the analytic solution will be given.

4.6.1 Charge Near a Planar Interface

One of the text book examples of solving Poisson equation with dielectrics is a single source charge sitting in front of a dielectric surface dividing two different dielectric regions of dielectric constant of ϵ_1 and ϵ_2 respectively. The situation is depicted in the Figure 4.5 (a), this is the situation for ICC algorithm. However, due to 3D periodicity in ICCP^{3}M boundary conditions, we must introduce another plane inside the simulation box, to prevent ill defined electric regions on the box boundary shown in Figure 4.5 (b).

The dielectric interface is populated with a rectangular grid of charges representing the polarization. For the algorithms we consider, there is no physical distinction among these charges and the source charges. We will investigate the polarization energy and the force acting on the source charge, which is called Dielectric Boundary Force (DBF) in the literature [42, 164].

We have investigated the values of the DBF on the charge, depending upon its location from the dielectric interface that divides, the region having a dielectric constant of ϵ_1 where charge sits and the the region having dielectric constant of ϵ_2. The dielectric interface is populated with induced charges, that each represent a square boundary element. The computation is repeated for four different corresponding location on the dielectric interface (out of plane), labeled (1) to (4), where the boundary charge sits on the location (3) as shown in Figure 4.6. This procedure gives us an opportunity to test the accuracy of the DBF depending upon source charge distance from the interface and corresponding out of plane locations on the interface.

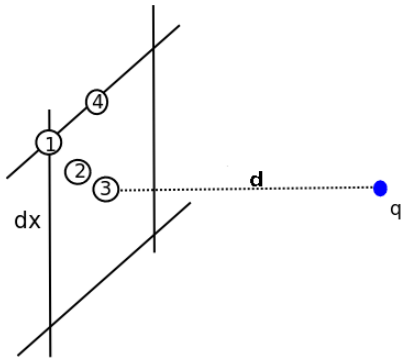


Figure 4.6: A single charge in front of a dielectric plane. The DBF on the charge is computed as a function of its distance from the interface and its corresponding location on the interface, labelled from (1) to (4) where induced charge align with the location (3).

ICC formulation for a Planar Interface

Let us denote the surface charge density for M grid points (induced charges) with 2D periodicity by h_α, with $\alpha = 1, ..., M$. In the simulations we used reduced the Bjerrum length l^* [1] of water scaled with ion diameter of 3 Å, which is equal to 2.321 at $T = 300$. N source charges have coordinates (x_i, y_i, z_i), with $i = 1, ..., N$, which form the subset of grid points (induced charges) $(x_\alpha, y_\alpha, z_\alpha)$. L_x and L_y be the plane dimensions.

$$h_\alpha = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} \cdot 2 \cdot l^* \sum_i^N \sum_n z_i \cdot \left[(x_\alpha - x_i - n \cdot L_x)^2 + (y_\alpha - y_i - n \cdot L_y)^2 + z_i^2 \right]^{-3/2}, \quad (4.35)$$

where n is an integer that specifies the periodic image box in both x and y directions. The electrostatic potential energy $\phi(r_p)$ of a system at the point p, where $r_p = (x_p, y_p, z_p)$ and a_α is the grid cell surface area has a form

$$\phi(r_p) = \frac{1}{4 \cdot \pi} \cdot \sum_\alpha^M \cdot \sum_{n_1} \sum_{n_2} h_\alpha a_\alpha \cdot [(x_\alpha - x_p - n_1 \cdot L_x)^2 + (y_\alpha - y_p - n_2 \cdot L_y)^2 + z_p^2]^{-1/2}, \quad (4.36)$$

where n_1 and n_2 are integers that specify the periodic image box in x and y directions. The total polarization electrostatic energy of the 2D-periodic system W is the sum of all electrostatic potentials at the source charge locations r_j, where $j = 1, ..., N$, namely:

$$W = \frac{1}{2} \cdot \sum_j^N \phi(r_j). \quad (4.37)$$

The analytic expression for the same observable can be presented in the form

$$W^{analytic} = \frac{1}{4} \cdot \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} \cdot \frac{1}{d} \cdot l^*, \quad (4.38)$$

where the distance d is shown in Figure 4.6.

[1] The Bjerrum length is defined as length where two charges feel 1 $k_B T$ of interaction energy.

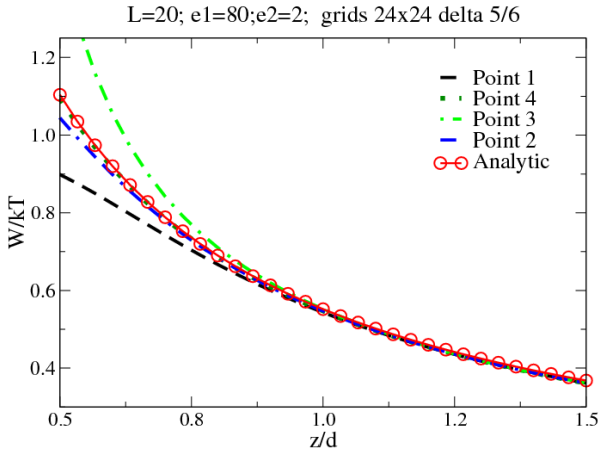


Figure 4.7: A single charge in front of a dielectric plane. Polarization energy W versus the distance of charge from the dielectric plane z, scaled with ion diameter d computed with ICC. Different curves represents corresponding locations on the planar interface. L=20 stands for the square frame length where induced charges placed, dielectric constants of $\epsilon_1 = 80$ where source charge sits and $\epsilon_2 = 2.0$, 576 grids (induced charges) on the dielectric plane placed equidistantly along x and y axis 24 by 24 within the square box frame of length L and delta equal to $5/6$ grid spacing ion diameter ratio.

In Figure 4.7 the comparison of ICC results for different source charge locations to the analytic solution is summarized. Here we plot the polarization energy versus the scaled distance from the planar dielectric surface. It is seen, that the closer the charge is to the dielectric surface the higher is the deviation of ICC curves compared to the analytical solution. But this effect is less pronounced if the source charge is located at the out of plane point (4).

Moreover, the effect of grid size on the dielectric plane is investigated for different values. The summary of polarization energy error, that is the qualitative deviation of ICC from the analytical result, this situation is given in Figure 4.8. Naturally, the smaller the grid size is the smaller deviation from the true solution is obtained.

We can compute the reaction force, DBF, $\mathbf{F}_{DBF}$, which is a force a unit charge feels at the location r_p due to polarization potential $\phi(r_p)$:

$$\mathbf{F}_{DBF}(r_p) = -\nabla\phi(r_p). \tag{4.39}$$

In component notation $\beta \in \{x, y, z\}$ the previous expression has the form.

$$\begin{aligned} \mathbf{F}^{\beta}_{DBF}(x_p, y_p, z_p) &= \frac{1}{4\cdot\pi}\cdot\sum_{\alpha}^{M}\sum_{n} h_\alpha a_\alpha \cdot (\beta_p - \beta_\alpha - n\cdot L_x)\cdot\Omega, \\ \Omega &= \left[(x_p - x_\alpha - n\cdot L_x)^2 + (y_\alpha - y_p - n_2\cdot L_y)^2 + z_p^2\right]^{-3/2}. \end{aligned} \tag{4.40}$$

Similarly, analytical expression for the force acting on the source charge due to surface charges can be written as follows

$$\mathbf{F}^{analytic}_{DBF} = \frac{1}{4}\cdot\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}\cdot\frac{1}{d^2}\cdot l^*\hat{\mathbf{z}}. \tag{4.41}$$

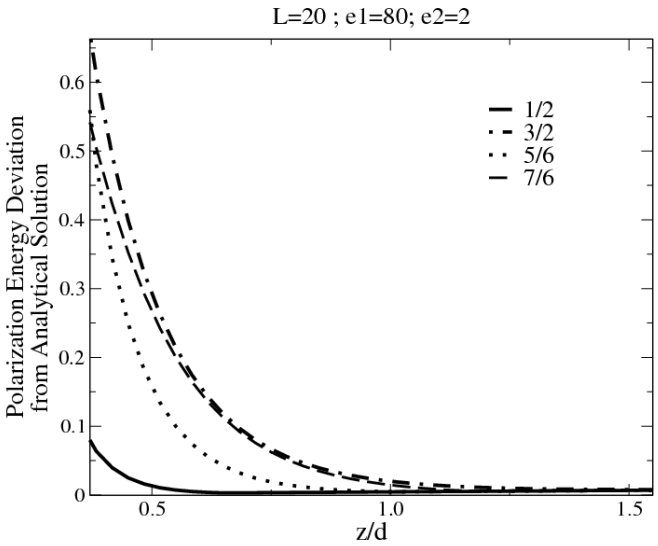


Figure 4.8: The error introduced into electrostatic energy due to discrete elements having different grid size ion diameter ratio of $\{1/2, 3/2, 5/6, 7/6\}$ as a function of scaled source charge distance from the dielectric interface. Computed with ICC.

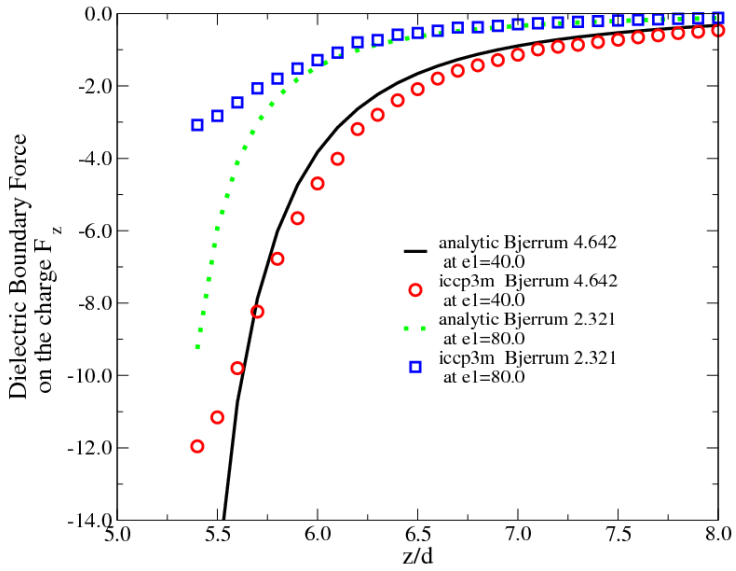


Figure 4.9: A single charge in front of a dielectric plane, where the total force on the charge is computed by ICCP^{3}M where the situation is the one depicted in Figure 4.5 (b). Plotted is the DBF versus distance. Different values of the dielectric constant, $\epsilon_1 = 40, 80$ is tested, with the corresponding Bjerrum length settings, $l^* = 4.642, 2.321$. Results is compared with analytical curves by using Equation (4.41).

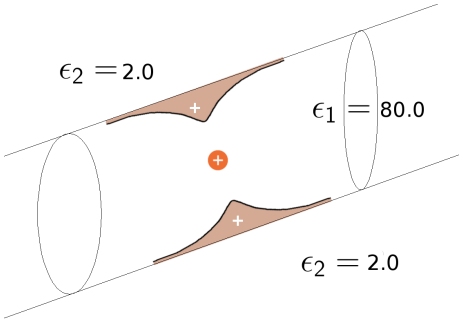


Figure 4.10: A single charge inside an infinite dielectric cylinder of dielectric constant ϵ_1 embedded in a bulk of another dielectric constant ϵ_2.

We have used ICCP^{3}M for computing the $\mathbf{F}_{DBF}$. Even though the ICCP^{3}M contains the forces from periodic images and the artificial symmetric interface as shown in Figure 4.5 (b), we obtain force profiles close to the analytic solution. The results are plotted in Figure 4.9 for two different dielectric constants, namely $\epsilon_1 = 40$ and $\epsilon_1 = 80$, corresponding to l^* scaled Bjerrum lengths 4.642 and 2.321 respectively, of the domain where the source charge sits. A fixed dielectric constant of $\epsilon_2 = 2.0$ for the other domain is considered for all cases. The important point is that corresponding Bjerrum length should be properly assigned to match the dielectric constant ϵ_1. The main observation is that, the higher the dielectric constant is in the region where the source charges sit, the higher the deviations is in the distances when source charges are near to the interface. This is due to the fact that, the electric field close to the interface is more distorted.

4.6.2 Charge inside an Infinite Cylinder

A single charge inside the infinite dielectric cylinder, with dielectric constant ϵ_1, which is surrounded by a dielectric media with dielectric constant ϵ_2 is depicted in Figure 4.10. Here we apply both ICC and ICCP^{3}M algorithms for comparison. Specially we focus on obtaining the induced surface charge density.

The elements of matrix A, A_{jk} first presented in Equation (4.27), which denote the derivative of Green function [48, 58], in the case of an infinite cylinder in cylindrical coordinates [179] are

$$\begin{aligned} A_{jk} &= \alpha_{jk}\frac{1}{4\pi^2}\sum_{m=-\infty}^{m=\infty}\exp(im\Delta\phi)\int_0^\infty dk\cos(-k\Delta z)k[I_m(kR)K_m^{'}(kR)+I_m^{'}(kR)K_m(kR)], \\ \alpha_{jk} &= \begin{cases} -(\epsilon_1-\epsilon_2), & \text{if } j\neq k, \\ (\epsilon_1+\epsilon_2)/2-(\epsilon_1-\epsilon_2), & \text{if } j=k, \end{cases} \end{aligned} \tag{4.42}$$

where I_m and K_m are the modified Bessel functions of the first and the second kind respectively. R is the channel radius, Δz and $\Delta\phi$ are longitudinal and azimuthal coordinate differences between two grid points (induced charges). This Green function for an infinite cylinder [48] has been used and tested for ICC computations. The expression for the c vector can be written as follows, that appear in Equation (4.27),

$$c_j = (\epsilon_1-\epsilon_2)\frac{q}{2\pi^2}\sum_{m=-\infty}^{m=\infty}\int_0^\infty dk\cos(k\Delta z)kI_m(kd)K_m^{'}(kR), \tag{4.43}$$

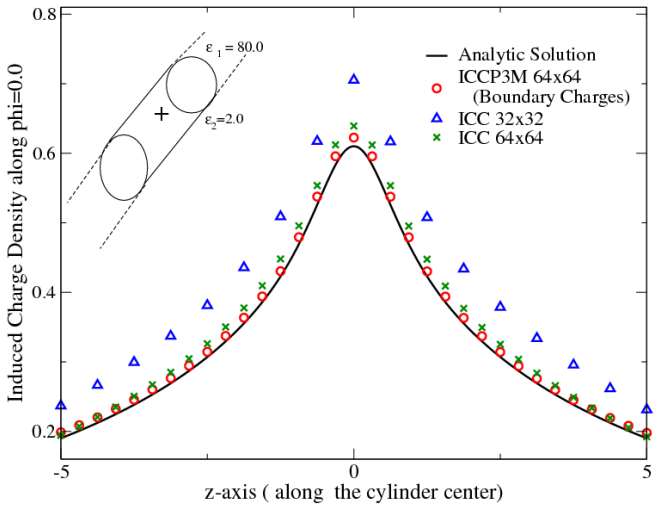


Figure 4.11: Surface charge density along the cylinder axis at a fixed azimuthal angle , due to a source charge sitting at the center computed with different grid schemes, NxM, where N is the number of charges along the cylinder axis z and M is the number of charges in the azimuthal direction. ICC with different grid schemes, ICCP^{3}M and analytic results are compared with setting $\epsilon_1 = 80.0$ and $\epsilon_1 = 2.0$ as it is depicted in Figure 4.10.

where d is the distance of the charge from the central cylinder axis (z-axis) and k is a real constant.

For implementation convenience, we can use the following identities

$$\frac{1}{x} = I_m^{'}(x)K_m(x) - I_m^{'}(x)K_m(x); \tag{4.44}$$

$$I_m(x) = I_{-m}(x); \tag{4.45}$$

$$K_m(x) = K_{-m}(x); \tag{4.46}$$

$$I_m^{'}(x) = I_{m+1}(x) + \frac{m}{x}I_m(x); \tag{4.47}$$

$$K_m^{'}(x) = -K_{m+1}(x) + \frac{m}{x}K_m(x), \tag{4.48}$$

where m is a positive constant.

The induced charge density on the dielectric interface obtained by using the ICC matrix formulation are summarized above. The boundary elements on the cylinder are placed equidistantly, NxM, being the number of total surface charges (located at the center of boundary elements), where N is the number of charges along the cylinder axis and M is the number of charges in the azimuthal direction. The same computation is repeated with the ICCP^{3}M in periodic geometries. The overall resulting polarization charge density obtained by two different methods is compared with the analytic solution, shown in the Figure 4.11, inside the cylindrical interface dielectric constant is $\epsilon_1 = 80.0$ and outside $\epsilon_1 = 2.0$. The analytic solution has been shown to be [48]:

$$\sigma(z,\phi) = l^* \sum_{m=-\infty}^{\infty} \exp(im\phi) \int_0^{\infty} dk A_m cos(kz), \tag{4.49}$$

$$A_m(k) = \frac{(\kappa-1)}{2\pi^2 R} I_m(kd) K_m'(kR) \left[I_m'(kR)K_m(kR) - \kappa I_m(kR)K_m'(kR) \right]. \tag{4.50}$$

We have truncated the given sum at an arbitrary order for a given integer m where the result does not change much but oscillate around a stable value, where the parameter κ is the ratio of dielectric constants ϵ_2/ϵ_1. Both methods yield comperable accuracies, even though they are not formally equivalent, while ICCP^{3}M treats the system in full periodicity in three dimensions. However ICC algorithm requires very large memory storage and this slows down the algorithm considerably compared to ICCP^{3}M . This fact renders ICC algorithm infeasable for large systems.

4.6.3 Charge Near a Dielectric Sphere

Here we consider a point charge placed at a distance d from the center of a dielectric sphere having dielectric constant ϵ_2, embedded in an environment of dielectric constant ϵ_1. The parameter κ is the ratio of dielectric constants ϵ_2/ϵ_1. The polarization surface charge density can be expanded in terms of the spherical harmonics, namely Legendre polynomials $P_l(\cos\theta)$ where θ is the polar angle in a spherical coordinate system. The expression of the charge density, found analytically in works [56, 176], has the following form

$$\sigma(\theta) = \sum_{l=0}^{\infty} A_l P_l \cos(\theta),$$

where coefficients A_l are given by

$$A_l = \frac{(1-\kappa)qR^{l-1}l(2l+1)}{4\pi\varepsilon_0 d^{l+1}(l(1+\kappa)+1)}.$$

In the present work we use $\kappa = 2$ being the ratio of dielectric constant and where R is the radius of the sphere. Again, we have obtained good agreement between the induced charge densities along a polar angle with ICCP^{3}M at a fixed azimuthal angle ϕ, the compared to the analytical results as shown in Figure 4.12.

We observe in Figure 4.12, if the source charges are placed further apart from the dielectric interface, small deviations from the analytic solution appear. However, this is due to the fact that we work in periodic boundary conditions to solve the algebraic problem for the induced charge density. Hence, when a source charge is sufficiently far away from the dielectric interface the effect of periodic image charges appear, even though the simulation box were 250 times larger then the radius of the spherical interface. A similar deviation from the analytic solution is seen when the source charge is close to dielectric interface, for example at $d = 1.2$. This is due to the fact that discretization effects on the dielectric boundary dominate.

4.6.4 Two Charges inside an Infinite Cylinder: Optimization

As a further test to find the accurate induced charge density we placed a pair of opposite charges a distance apart on the central axis of infinite dielectric cylinder in a fully periodic system in three dimensions. Due to symmetry reasons when the separation distance between two charges is determined to be half of the box length we expect to get zero total force on the charges, which we have confirmed.

The resulting force profile is depicted in Figure 4.13 with different Bjerrum lengths that correspond to different inner dielectric constants. Note, that the force has only a net contribution along the cylinder central axis. The forces reported in the figure represent only the force due to the induced charges, a periodic DBF.

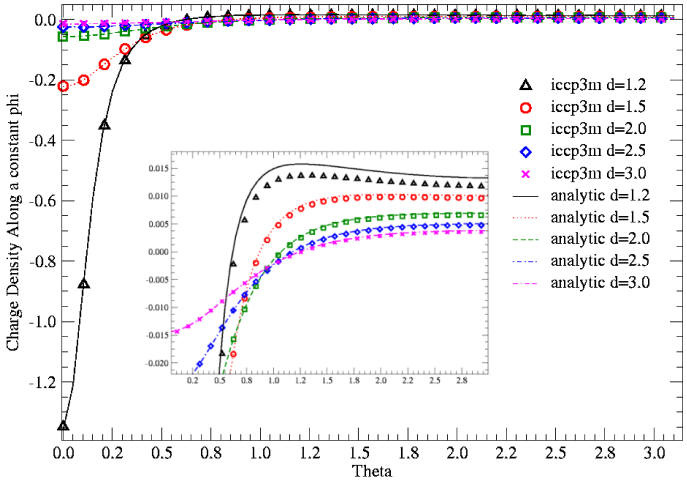


Figure 4.12: Plot of the induced charge distribution on the spherical interface at a constant azimuthal angle ϕ, and varying polar angle θ. Induced charges occur due to a source charge located at d distance apart from the sphere. ICCP^{3}M reproduce the example solved via density functional method [48].

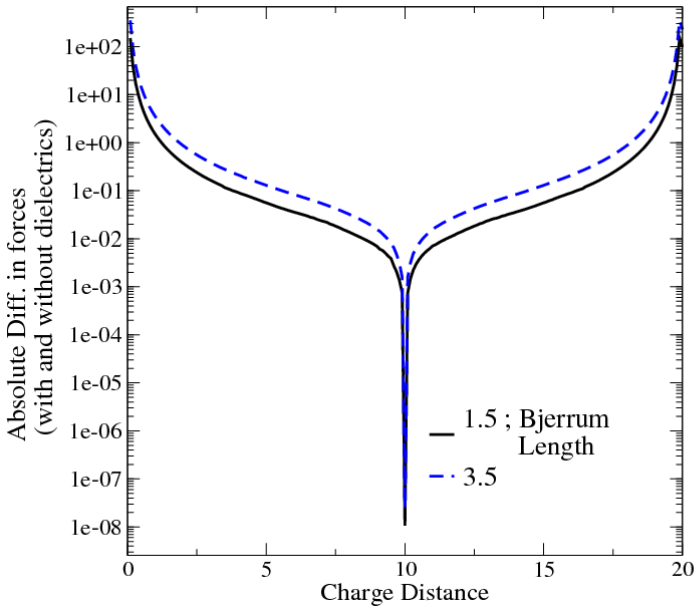


Figure 4.13: The force profile on the two opposite charges a distance apart on the central axis of a dielectric cylinder for different Bjerrum lengths that corresponds to different inner cylinder dielectrics. Forces are zero when two charges are separated by a distance of half the box length due to symmetry. It seen that forces are way below P^{3}M accuracy hence it is safely considered to be zero.

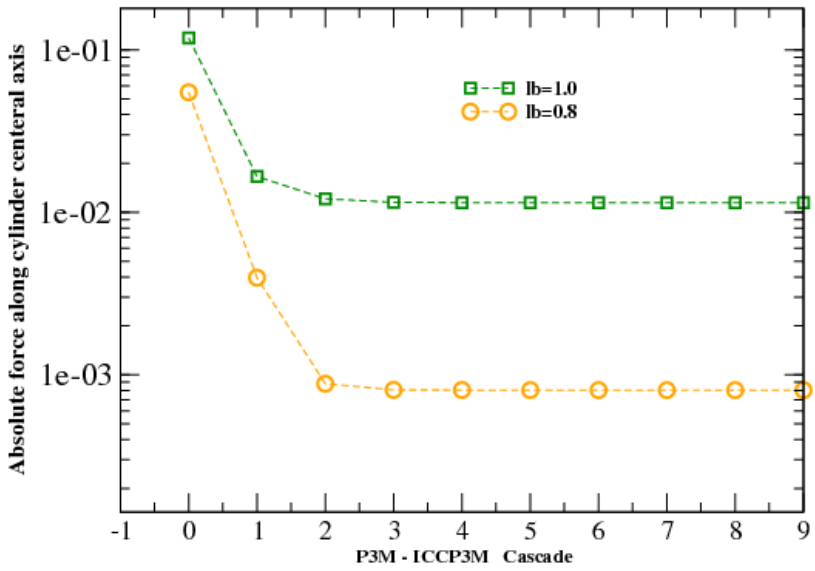


Figure 4.14: Two charges a distance apart on the central axis of a dielectric cylinder: Optimization with P^{3}M - ICCP^{3}M cascades. Notice that cascade labeling starts from zero.

During this example, we have noticed that applying ICCP^{3}M procedure a couple of times in sequence, make the charge values change significantly. That is to say P^{3}M mesh size and charge assignment parameters adjusted once again after ICCP^{3}M iteration, namely tuned. Then, ICCP^{3}M routine is called once again, until the induced charges or the DBF reach to a stable value. The primary reason behind this behaviour is the fact that the parameters of P^{3}M directly related to charge values present in the system. If the charge values in the system is changing due to ICCP^{3}M iteration, as a result P^{3}M parameters are no longer represents the optimized values, hence must be tuned again. The effect of this procedure is shown in Figure 4.14 for two different Bjerrum lengths that corresponds to two different dielectric constants for the cylinder. Note that in the figure we only report total force on the charge, that includes source-source interactions.

This kind of procedure can be thought as a cascade of P^{3}M parameter optimization - ICCP^{3}M iteration cycle. We have given labels to this cycles in Figure 4.14 staring from zero. For this particular geometry and system the total force on the charge is reaching a stable value after a couple of cascades.

We have shown how to optimize the ICCP^{3}M . This optimization is crucially needed since the optimized forces differ almost one order of magnitude from the non-optimized ones, as can be expected by looking at the label 0 and label 4 cascade forces given in Figure 4.14.

4.7 ICC and ICCP^{3}M Performance Comparison

ICCP^{3}M has advantages in comparison to other similar methods developed in the frame work of the ICC method. The first advantage is that it is computationally faster than the ICC method. Unlike the other ICC methods mentioned earlier, it shows an improved scaling of $\mathcal{O}(M \ln M)$ rather than a $\mathcal{O}(M^2)$ scaling, M being the number of induced charges when the number of source

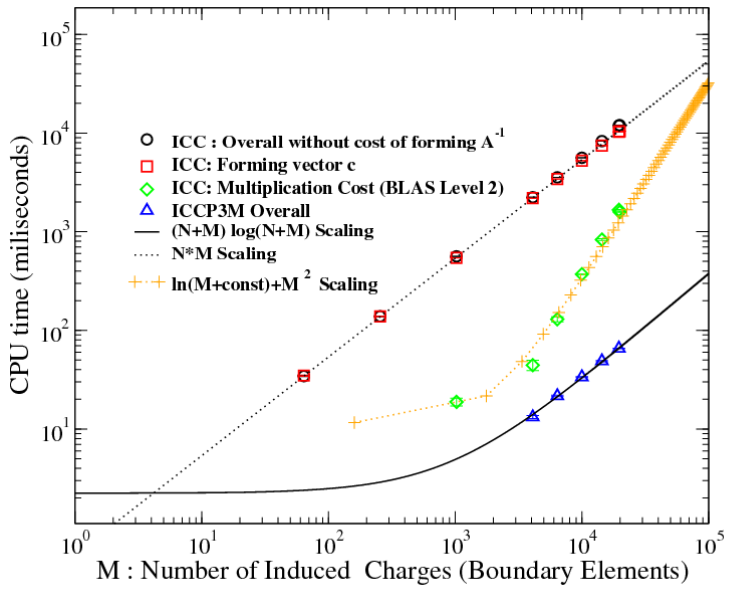


Figure 4.15: A performance comparison between ICC and ICCP^{3}M for a dielectric cylinder. The scaling of different computation chunks of ICC algorithm is reported. For example, computation of vector c (source charge - boundary charge interactions) and the multiplication cost $h = A^{-1}c$. In order to be fair to ICC method, timing of forming the matrix A and its inverse is ignored, essentially only multiplication cost remains. Even the optimized, with BLAS, vector-matrix multiplication cost scales as $\mathcal{O}(M^2)$, where ICCP^{3}M overall induced charge computation procedure scales like $\mathcal{O}(MlogM)$.

charges is fixed.

The scaling of computational cost against the number of induced charges M, on the dielectric interface is computed and compared for the ICC method and the ICCP^{3}M in Figure 4.15. The computation is performed for a fixed number of source charges N, located inside a dielectric cylinder, and timings are averaged over 100 runs. The scaling of different computational parts of the ICC algorithm is reported. For example, the computation of vector c (source charge - boundary charge interactions) and the multiplication cost $h = A^{-1}c$. In order to be fair to the ICC method, we did not take into account the time to form the matrix A and its inverse, in all ICC comparison cases, essentially making the overall ICC costs only a multiplication cost shown in Figure 4.15. For ICCP^{3}M , the timing is taken one of full iteration scheme with a convergence criterion of 10^{-4}. We also reported the multiplication cost of $A^{-1}c$ with an optimized library BLAS. We see that even this multiplication cost is scaling as $\mathcal{O}(M^2)$. However, overall ICCP^{3}M scaling is seen as $\mathcal{O}(MN\log(MN))$.

The second advantage is that we can work with genuine periodic boundary conditions. In other ICC methods, the periodic boundaries conditions are not used full effectively. In the Fast Multipoles method the implementation of periodic boundary conditions is complicated. In the ICC type methods, these boundary conditions are approximated using the nearest neighbor interaction that is not quite good since the electrostatic interactions are long range interactions and they should not be cut-off.

The third advantage is that ICCP^{3}M iterations converge very fast. As explained above, when

we use as an initial distribution the values of the induced charges from the previous step of the molecular dynamic simulation, the results converge quickly as we already start with values close to the correct ones. Usually in this case study, iterative procedure converges in one or two iterations.

The fourth advantage is that the method requires only a very modest amount of memory. In the ICC method, because of an explicit construction of matrix A, this memory requirement happens to be a huge bottleneck, or even render the computation infeasible. For example, if we need five thousand points to decorate the interface, we would require enough memory to hold a matrix of size 5000 times 5000, which amount to $\sim$ 400 MB, where in case of ICCP^{3}M this would only yield to an amount of order of several MB, depending upon P^{3}M mesh size.

4.8 Summary

The problem of finding correct and accurate electrostatic forces on the charges embedded in a dielectric medium that has an interface with an other dielectric medium has been discussed. The provided general introduction showed the importance of the problem and motivated the methods put forward in this chapter.

We have presented the surface charge polarization conceptually and its quantitative description both as solving a functional minimization and a system of linear equations (ICC algorithm), on the induced charge density. The introduction of an iterative solution on the system of linear equations for the induces charge densities produced a computationally much more favorable algorithm due to fact that manipulating and storing a matrix is not feasible for large systems. Furthermore utilization of the P^{3}M in the force computation of the iterative procedure gives rise to a novel hybrid technique ICCP^{3}M .

We validated our implementation on simple geometries where analytical solutions were available. Results are shown for the correctness of both implementations and algorithms achieved in describing induced charges and Dielectric Boundary Forces (DBF). Moreover we have optimized the algorithm by introducing a further cascading procedure. Supported by our test example, we have shown the crucial effect of this optimization on the DBF values.

Further computational performance measurements were carried out in finding induced surface charge densities against an increasing number of surface charges. Timing for different parts of the ICC algorithm were provided, including vector-matrix multiplication costs and forming the source-induced charge interaction vector. We have shown that the ICCP^{3}M overall procedure out performs even the multiplication cost. Moreover, the memory requirements of the ICCP^{3}M grows almost linearly with the increasing number of surface charges, $\mathcal{O}(M)$, while in the case of ICC, it increases with $\mathcal{O}(M^2)$ in the best case. We have concluded that the ICCP^{3}M is a more superior algorithm then the ICC algorithm.

5 Model DNA and Effect of Dielectrics in Translocation

In this chapter we address the effect of dielectrics on macromolecular transport through nanopores, focusing on a coarse-grained model of stiff DNA and its translocation. In Section 5.1, a general outlook on polymer translocation and related problems in statistical physics is given. A newly developed coarse grained model for DNA is presented in Section 5.2, as well as an overview on other relevant DNA models. Counterion condensation around DNA is studied for a rod-like conformation. Investigation of the effect of the channel geometry on the surface polarization charge distribution is discussed in Section 5.3 for a single charge located at the centre of a finite cylindrical channel in a fully periodic simulation box. The translocation of stiff DNA through a nanopore is studied in Section 5.4 by computing its potential of mean force with and without the inclusion of a dielectric discontinuity.

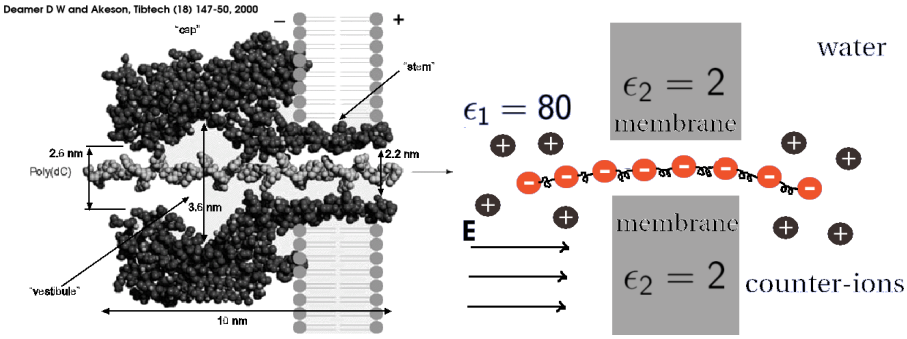
5.1 Macromolecular Transport: An Outlook

Single Molecule Experiments (SMEs) are employed to understand biological processes at the molecular scale. Of special interest is the transport of macromolecules [7] through natural nanopores, e.g., nucleic acid or proteins [180, 181, 104, 103, 5, 113, 109, 114], and, through manufactured solid-state nanopores [6, 31, 32, 120]. These studies can open a new dimension for our understanding of molecular properties, e.g., the direct measurement of DNA effective charge, and are the first step towards single molecule sequencing techniques [33, 34, 35].

Ion channels importance and function in the living organism is very well known,and there is a large amount of experimental, theoretical and computer modelling studies [91, 98, 182]. Even though there is a vast amount of information and many mechanisms are known, there are still open fundamental questions, such as ion channels selective permeability.

Ionic current measurements are possible, and current-voltage characteristics of different channels have been investigated in depth experimentally, theoretically and computationally [183, 184, 185]. Moreover, the gating mechanism of ion channels appear also as an open research problem [186], and the electrostatic interactions seem to play an a vital role in this mechanism [187]. From the computational point of view, continuum theories (such as Poisson-Nernst Planck Equations), stochastic dynamics, and molecular dynamics techniques have been employed to investigate these questions [188].

In the same context, biopolymer translocation through narrow nanopores has been studied in depth experimentally, theoretically and with computer simulations [5]. Landmark experiments performed by Kasionowicz et al [94] and Meller et al [102] for voltage assisted translocation of polyelectrolytes through narrow nanopores have stimulated theoretical and simulation activities in the last decade [104, 103, 107, 108, 5, 109, 113, 109, 114, 189, 190, 191, 192, 111]. In these experiments a polyelectrolyte is driven by an external electric field into the pore blocks ionic current. The ionic current characteristics can be analyzed when the chain is inside the channel, in



(a) Crystallographic structure of single strand (ss) DNA and alpha-hemolysin channel. The ss-DNA is driven trough the narrow channel with the assitance of voltage difference between membrane sites.

(b) Corresponding typical coarse grained model of the ss-DNA and the alpha-hemolysin channel system, that is suitable for coarse grained dynamical simulations of voltage driven translocation.

Figure 5.1: Coarsed graining for polymer translocation.

order to infer its sequence. Experiments in this direction were not limited to biological pores, but were also applied to solid-state pores [6].

A large number of studies showed the dependence of translocation time on different system characteristics, such as the strength of the external field and the polymer chain length. However, the effect of dielectric domains has been studied by only few groups, e.g. simulation of flexible DNA inside the nanopore using coarse-grained molecular dynamics [113, 114], still, the treatment of the dielectric jump between solute and the membrane was just phenomenological and not formally exact.

Many of the simulation studies in the context of macromolecular transport employ coarse-grained approaches [36] due to the spatio-temporal scales inherent to the systems. One of the steps in realizing a coarse-grained model for such problems is to choose a suitable representation of the solvent, especially in presence of dielectric interfaces. In Chapter (4) we already discussed in depth the mathematics and algorithmics that can be employed to treat dielectrics in coarse-grained simulations[42, 164, 44], proposing a fast and efficient algorithm to compute dielectric boundary forces (DBF). Another important step, which is addressed in the next sections, consists in modelling the solute. The standard approach, which we will also employ, is to use bead-spring models [193]. A typical situation that depicts the coarse graining for a biopolymer translocating through a protein membrane channel is presented in Figure 5.1.

5.2 A Simple Model for DNA

Due to spatial and time scale restrictions of full atomistic simulations of soft matter and biophysical systems, coarse-grained models have become widespread tools [36, 37, 38]. This applies also to the study of DNA translocation through nanopores, that we want to address and, therefore, a suitable model of DNA is required.

The initial step in developing the coarse-grained model is to choose an appropriate length scale and to decide which degrees of freedom are necessary to recover available experimental observations. Not every property can be reproduced by a coarse-grained model though.

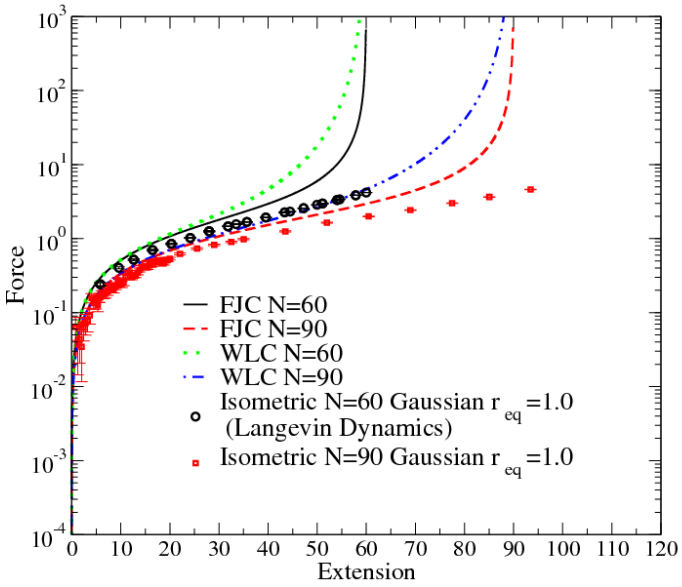


Figure 5.2: The comparison of isometric Gaussian stretching simulation data discussed in Chapter 3 against WLC and FJC analytical predictions.

Force-extension curves obtained via SMEs can be reproduced by a model that does not carry specific information of the molecular structure and it is in general sufficient to represent the chain as a generic flexible string. However, in some situations one would need a more specific model, for example if one wants to model DNA denaturation or various thermally activated processes. An intermediate scale model has been therefore developed [194, 195]. In the intermediate scale one can deal with physics not that far from the one reproduced by full atomistic models as we will see in section 5.2.3.

A coarse-grained polyelectrolytes model of DNA has been constructed for counter-ion condensation studies [39, 40]. In these cases, simple bead-spring model with correct DNA geometrical and electrical characteristics is sufficient to capture the essential physics.

5.2.1 Worm Like Chain Model

As we have mentioned in Chapter (3), information from stretching biomolecules is quantitatively accessible as force-extension curves. The simplest analytical description of a polymer has been formulated in terms of the Gaussian chain. However, a more realistic model of DNA stretching [18], that can reproduce experimental force-extension curves is the Kratky-Prod model, the special case of which is also known as Worm Like Chain (WLC). In WLC the polymer is modeled as a flexible string.

The energy, E, of a streched DNA molecule is given in the WLC model by a line integral of two terms [196]. The first term, e_b, describes the resistance of the chain to bending and is proportional to the inverse square of the radius of curvature, R, which is related to the derivative of the tangent vector $\frac{d\mathbf{t}}{ds}$ as $R = 1/|\frac{d\mathbf{t}}{ds}|$. One gets then

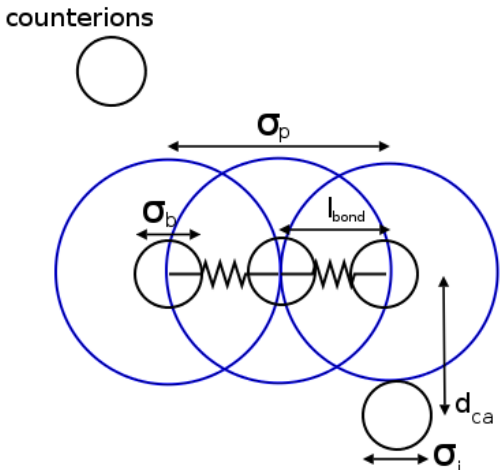


Figure 5.3: The sketch of the Antypov coarsed grain DNA model. The length scales specified in the figure stands for, a is the polymer bead (mimics a base-pair) diameter $\sigma_p = 15.72$, b is the polymer FENE bead diameter $\sigma_b = 1.87$, c is the FENE bond length $l_{bond} = 1.1\sigma_b$, d is the distance of closest approach to polymer bead which is 9.98 for coions and counterions, e is the coion and the counterion diameter, both $\sigma_i = 4.25$. In the simulations LJ length scale is taken to be 1.7, which is the charge separation between two DNA base-pairs, that is extracted from DNA line charge density.

$$e_b = \frac{A}{2}\left|\frac{d\mathbf{t}(s)}{ds}\right|^2 = \frac{A}{2R^2}. \tag{5.1}$$

The second term gives the streching energy resulting from the application of the force F, to the end of the molecular chain. Taking the force along the z axis, one arrives to the final espression for E_{WLC}:

$$E_{WLC} = \int_0^{L_0} ds(e_b - Fcos(\theta(s))), \tag{5.2}$$

where $\cos\theta(s)$ is the angle between $\mathbf{t}(s)$ and the z axis, and L_0 is the contour length of the molecule. Experimental data are usually analyzed in terms of the persistence length, L_p, which is related to the quantity A by $L_p = A\beta$, where $\beta = 1/k_BT$.

One can write a partition function of a polymer by using the energy E_{WLC} and derive an explicit formula for force laws. However, it is not straightforward to solve the resulting Schrödinger-like differential equation [10]. Solutions to force laws are presented either as numerical approximations, [10] or under special analytical restrictions [18, 10]:

$$F_{WLC} = \frac{k_BT}{L_p}\left[\frac{1}{4(1-X/L_0)^2} - \frac{1}{4} + \frac{X}{L_0} + \sum_{i=2}^{i\leq 7} a_i\left(\frac{X}{L_0}\right)\right], \tag{5.3}$$

X is the extension and the coefficients values are $a_2 = -0.5164228$, $a_3 = -2.737418$, $a_4 = 16.07497$, $a_5 = -38.87607$, $a_6 = 39.49944$, $a_7 = -14.17718$ [10]. Although it is complicated

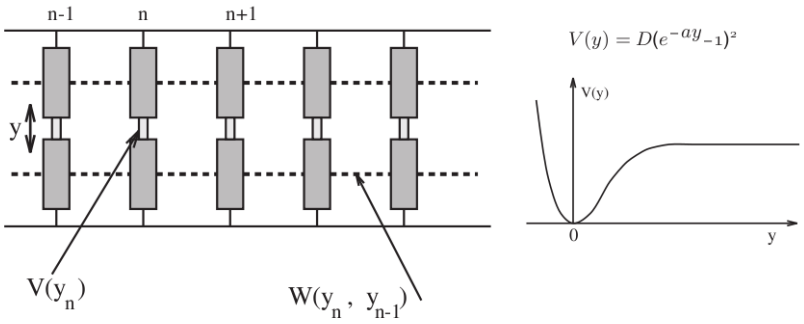


Figure 5.4: The Simple DNA model (Peyrard-Bishop-Dauxois Model) for molecular dynamics [195].

to mimic WLC in Langevin Dynamics with bond constraints, recently developed bead-spring like models [197] made the Langevin Dynamics study of the DNA stretching possible [198].

Another simple model, which assumes discrete chain segments with free orientations is the Freely Joint Chain (FJC) one, and the force law for this model is [10]

$$\frac{X_{FJC}}{L_0} = coth\left(\frac{2L_pF}{k_bT}\right) - \frac{k_BT}{2L_pF}.$$

The comparison of these force laws to the computer simulations of a Gaussian chain with $r_{eq} = 1.0$ is given in Figure 5.2. Even though theoretical curves are all close to simulation results, it is seen that the FJC model turns out to be the best approximation. The units $k_BT = 1.0$ and $L_p = 1.0$ have been used.

5.2.2 Peyrard-Bishop-Dauxois Model

In describing the phenomenology of DNA dynamics, such as thermal denaturation, models that captures the coarse-grained local structure sufficiently well were developed [194, 199, 200, 195] and further formulated for Langevin Dynamics [201].

The primary information that DNA stores is known as the genetic code which is realized by the sequence of DNA base-pairs. Hence, a model that takes into account the base-pair as a basic unit is suited to capture the essential physics in denaturation. The Peyrard-Bishop-Dauxois (PBD) model is one of the model depicted in Figure 5.4 [195].

The variable in the PBD model is the transverse stretching q_n of the hydrogen bond connecting two base pairs, labeled with the integer n, as shown in Figure 5.4. Negative values of q_n are interpreted as a compression of the bond. The Hamiltonian of the model system can be written as

$$H(p_n, q_n) = \frac{p_n^2}{2m} + W(q_n, q_{n-1}) + V(q_n), \tag{5.4}$$

$$p_n = m\frac{dq_n}{dt}, \tag{5.5}$$

where m is the reduced mass of the $n-th$ base and p_n is an effective momentum. Notice that bases are indistinguishable. There are two potentials representing the interactions between two

consecutive base-pair namely, $V(q_n)$ and $W(q_n, q_{n-1})$. They are approximated with harmonic and Morse type effective interactions as follows

$$W(q_n, q_{n-1}) = \frac{1}{2}K(y_n - y_{n-1})^2, \tag{5.6}$$

$$V(y_n) = D(exp(-ay_n) - 1)^2, \tag{5.7}$$

where D is the disassociation energy, a is a free parameter which sets the spatial scale of the potential, and K is the spring constant. The typical values suggested in reference [195] are: $D = 0.03$ eV , $a = 4.5$ Å^{-1} and $K = 0.06$ eV Å^{-2}. It is usual in coarse-grained models to use dimensionless expressions for basic quantities. This can also be achieved for the PBD model by transforming the PBD Hamiltonian into $H'(P_n, Y_n) = H(p_n, q_n)/D$. The canonical variables will therefore

$$Y_n = ay_n, \tag{5.8}$$

$$P_n = \frac{dY_n}{d\tau}. \tag{5.9}$$

Here $\tau = (Da^2/mt)^{1/2}$ is the reduced time, and $S = K/(Da^2) = 0.0976$. Hence, the dimensionless Hamiltonian has the form:

$$H'(P_n, Y_n) = \frac{1}{2}P_n^2 + \frac{1}{2}S(Y_n - Y_{n-1})^2 + (exp(-Y_n) - 1)^2. \tag{5.10}$$

As we have already mentioned this model can be efficiently used to study DNA denaturation, but has a resolution higher than the one needed for performing DNA stretching simulations or investigating DNA charge characteristics. This problems will be discussed below.

5.2.3 Coarsed Grain DNA: Antypov Model

In this model, an overlapping excluded volume interaction is introduced, among DNA backbone charge centers (a single base-pair charge) and ions. DNA backbone charges are separated in accordance to the DNA line charge density in solution. Counter ions or salt ions are explicitly represented. The overlapping interaction approach is employed to mimic the surface structure of DNA at the coarse-grain level, which is well suited for stiff polyelectrolytes, such as ds-DNA (double-stranded DNA). The base-pair charge centers are connected with Finitely Extensible Nonlinear Elastic (FENE) springs [193]. Excluded volume and FENE interactions are sketched in Figure 5.3. The geometrical and charge characteristics of DNA which have been used [40], are summarized in Table 5.1.

The Lennard-Jones interaction for ion-ion (counterion or coion combinations of r_{i-i} distance apart) is

$$U^{i-i}(r_{i-i}) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma_i}{r_{i-i}}\right)^{12} - (\frac{\sigma_i}{r_{i-i}})^6 + \frac{1}{4}\right] & r_{i-i} < r_{cut} = 2^{1/6}\sigma_i = 2.81, \sigma_i = 2.5, \\ 0 & r_{i-i} > r_{cut}, \end{cases}$$

for ion - polymer bead (r_{i-b} distance apart) is

$$U^{i-b}(r_{i-b}) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma_i}{r_{i-b}-r_s}\right)^{12} - \left(\frac{\sigma_i}{r_{i-b}-r_s}\right)^6 + \frac{1}{4}\right] & r_{i-b} < r_{cut} = r_s + 2^{1/6}\sigma_i = 6.18, r_s = 3.38, \\ 0 & r_{i-b} > r_{cut}, \end{cases}$$

Parameter	Symbol	Value	Value in LJ Units
Length Scale	σ	1.7000	1.0000
Ion Diameter	σ_i	4.2500	2.5000
Overlapping Bead Diameter	σ_p	15.7200	9.2470
FENE Bead Diameter	σ_b	1.5454	0.9090
Distance of Closest Approach	d_{ca}	9.9850	5.8735
LJ Shift (Translation Parameter)	r_s	5.7350	3.3735
Bjerrum Length	l_B	7.1400	4.2000
Line Charge Density	λ	$e_0/1.7$	e_0/σ
FENE Max Extension	R_{FENE}	3.0909	1.8181
FENE energy constant	k_{FENE}	7ϵ	7.0
FENE Bond Length	l_{bond}	1.7000	1.0
Bond Angle Equilibrium	θ_0	0	0
LJ Energy Parameter	ϵ	$0.25k_bT$	0.25
Bond angle energy constant	k_a	40ϵ	40
Temperature	T	$298^\circ K$	e/k_B
Box Length	$L = \left(\frac{4}{3}\pi\sigma_i^3 N_c c^{-1}\right)^{1/3}$		

Table 5.1: The System Parameters for the DNA model, length values are given in Å.

and for polymer non-overlapping bead - bead (r_{b-b} distance apart) is

$$U^{b-b}(r_{b-b}) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma_b}{r_{b-b}}\right)^{12} - \left(\frac{\sigma_b}{r_{b-b}}\right)^{6} + \frac{1}{4}\right] & r_{b-b} < r_b = 2^{1/6}\sigma_b = 1.02 \\ 0 & r_{b-b} > r_{cut}. \end{cases}$$

The FENE potential for polymer bead-bead model is

$$U^{\mathrm{FENE}}(r) = -\frac{1}{2}k_{\mathrm{FENE}}R_{\mathrm{FENE}}^2 \ln\left(1 - \frac{r^2}{R_{\mathrm{FENE}}^2}\right),$$

where $k_{\mathrm{FENE}} = 7.0$ and $R_{\mathrm{FENE}} = 1.82, \sigma_b = 0.91$. The bond angle potential for the polymer overlapping bead-bead bending interaction is

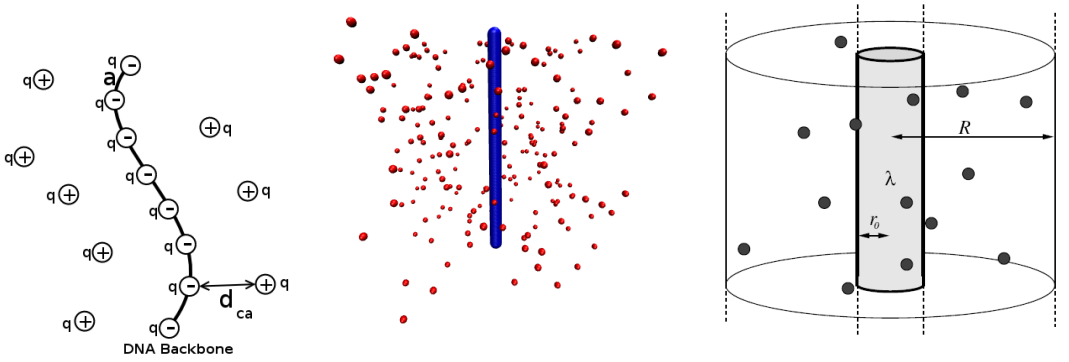
$$U^{angle}(\theta) = \frac{1}{2}k_a(\theta - \theta_0)^2,$$

where $k_a = 40$ and $\theta_0 = \pi$.

Notice that for a stiff DNA model, the potentials $U_{FENE}(r)$ and $U^{angle}(\theta)$ can be neglected while they do not change in the course of simulations. Next, we investigate the counterion condensation around a stiff DNA rod by coarse-grained Langevin Dynamics. The main purpose of this check is ensuring introducing overlapping backbone interactions against counterions do not introduce artifacts on the electrostatic properties of the model DNA.

5.2.4 Counterion Condensation

Strongly charged rod like polyelectrolytes tend to reduce their overall charge by attracting oppositely charged ions i.e. via counterion condensation. Counterion condensation for rigid polyelec-



(a) Sketch of the charge centres in model DNA with a flexable backbone and counterions. d_{ca} is the *distance of closest approach.*

(b) A snapshot of simulation of DNA rod and counterions. DNA rod is placed at the center of the simulation box subject to periodic boundary conditions.

(c) Sketch of the cell model [40]. An infinite rod of line charge density λ and radius r_0, surrounded by counterions in a cylinderical box of radius R.

Figure 5.5: DNA model with counterions, and a cell model.

trolytes has been understood well and studied in depth in the context of the cell model Poisson-Boltzmann formulation [40, 202]. There, an infinite charged rod inside an infinite cylindrical cell was studied together with counter ions.

The line charge density, λ, of DNA is used to determine the distance between two base-pair charges, which is taken to be 1.7 Å[40]. The rod charges are aligned on one axis in the simulation box as shown in Figure 5.7.

Let us summarize the major understanding of counterion condensation phenomena:

- Manning proposed that there is a clear distinction between a condensed and diffuse layer (Debye-Hückel) of counterions around a charge object, e.g. a charged rod.
- There exists an analytic solution of radial counterion distribution for no added salt case via salt-free cell model Poisson-Boltzmann (PB) description.
- The integrated radial counterion distribution, $P(r)$, which corresponds to the cumulative number of ions counted from the centre of the charged rod up to a distance r, contains quantitative information about the distinction between two charge distribution layers. So, there is an inflection point on this curve that locates the end of condensed layer and the beginning of the diffuse layer.
- The Manning parameter ξ is defined as the basic measure of electrostatic interaction strength

$$\xi = \lambda l_B / e,$$

where e is the elementary charge and $l_B = e^2(4\pi\epsilon k_B T)^{-1}$ is the Bjerrum length.

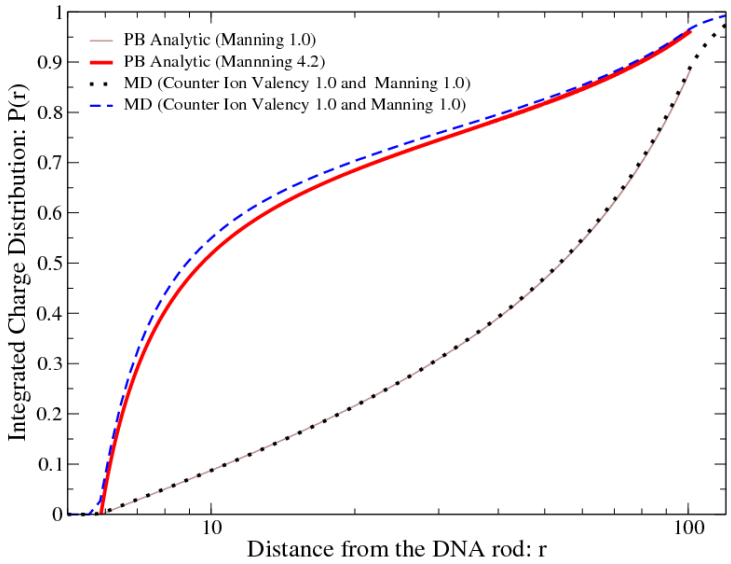


Figure 5.6: Coarse-Grained DNA Poisson-Boltzmann and MD comparison of integrated charge distribution with two Manning parameters 1.0 (lower curve and dots) and 4.2 (upper curve and dots). System with multivalent counterions has also been simulated.

- The salt-free cell model Poisson-Boltzmann (PB) provides [202] the following expression for the integrated radial counterion distribution

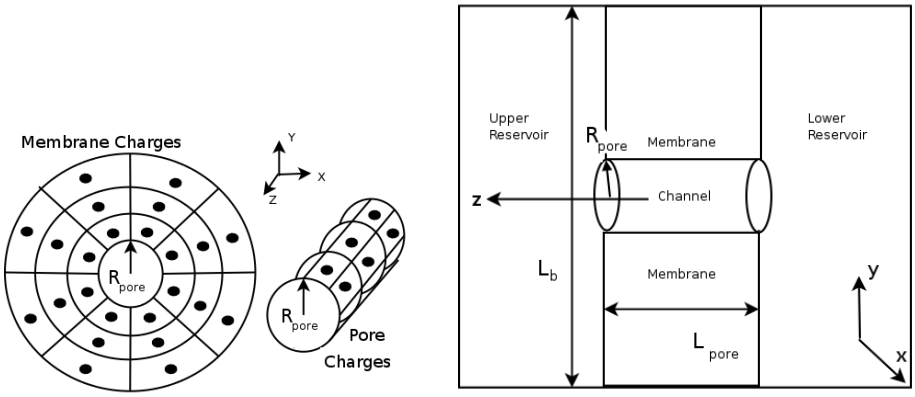
$$P(r) = \left(1 - \frac{1}{\xi}\right) + \frac{\gamma}{\xi} \tan\left[\gamma \ln\left(\frac{r}{R_M}\right)\right],$$

  where γ is a free parameter and R_M is the Manning radius, both determined via the boundary conditions.

We have simulated our coarse grained DNA model to understand counterion condensation around it with Langevin Dynamics and have compared the results to the cell model PB predictions for different Manning parameters and counterion valency. The overall comparison between Langevin Dynamics and the PB prediction is given in Figure 5.6. The simulations have been performed using two different Manning parameters, namely, 1.0 and 4.2, employing monovalent counterions. Furthermore, condensation has been investigated in cases of divalent and trivalent counterions.

The major observation is that PB failed to reproduce the simulated integrated counterion distribution when Manning Parameter is 4.2 and counterions are multivalent. However, in case of Manning parameter equals to 1.0 with monovelent ions the simulation results agree with the PB predictions. This test clarifies that the overlapping excluded interactions in our model do not generate any artifacts from the electrostatics point of view.

Before applying Antypov model to the problem of DNA translocation through a finite channel we will first discuss the effect of a channel-membrane geometry on the polarization charge density.



(a) Setting of polarization surface charges on the center of each boundary element for the membrane side (*membrane charges*) and for the cylindrical pore side (*pore charges*).

(b) A channel geometry mimics an ion channel in a simulation box, that is periodically replicated in three spatial dimensions.

Figure 5.7: Dielectric boundary charges and a channel geometry.

5.3 Effect of the Channel Geometry On the Induced Charge Density

The striking effect of the dielectric boundary force in ion channels has been pointed out in Chapter 4 and the approach to compute this force for arbitrarily shaped domains has been presented.

To understand the effect of the finite channel geometry on the polarization charge we have used ICCP^{3}M for a single charge located at the centre of the finite cylindrical channel in simulation box, that is fully periodic, see Figure 5.7.

The parameters to be set are: simulation box length L_b, cylindrical pore length L_{pore}, radius of the pore R_{pore}, the number of boundary elements on the cylindrical pore located along the $z-$direction M, and the number of elements at each azimuthal direction M. Meaning rings of charges form the cylindrical pore. Both sides of the membrane are covered by M concentric rings of charges. There are M charges in each concentric layer. The boundary element configuration is shown in Figure 5.7.

The distance between the centres of two consecutive boundary elements along $z-$direction of the cylindrical pore is $\Delta_{pore} = L_{pore}/M$. The distance between two concentric layers on the membrane side is $\Delta_{membrane} = (L_b - 2R_{pore})/M$.

The set of geometrical and spatial parameters that could effect the ICCP^{3}M procedure in finding the induced charge density (distribution) on the boundary due to the prence of a single charge situated at the box/pore centre can be counted as follows: $L_b, L_{pore}, r, \epsilon_1, \epsilon_2$, Δ_{pore},$\Delta_{membrane}$ and M. We vary the pore length L_{pore} and the radius R_{pore} in separate sets of calculations to understand their effect on the resulting induced charge distribution, keeping other parameters fixed. Note, that variying these parameters involves the changes of Δ_{pore} and $\Delta_{membrane}$. The set of variables choosen allow to keep the number of charges M fixed. One has to be also accurate when choosing the ratio between the mesh size of P^{3}M and distances Δ_{pore} and $\Delta_{membrane}$. The mesh size of P^{3}M should remain significantly smaller than any of Δ_{pore} and $\Delta_{membrane}$.

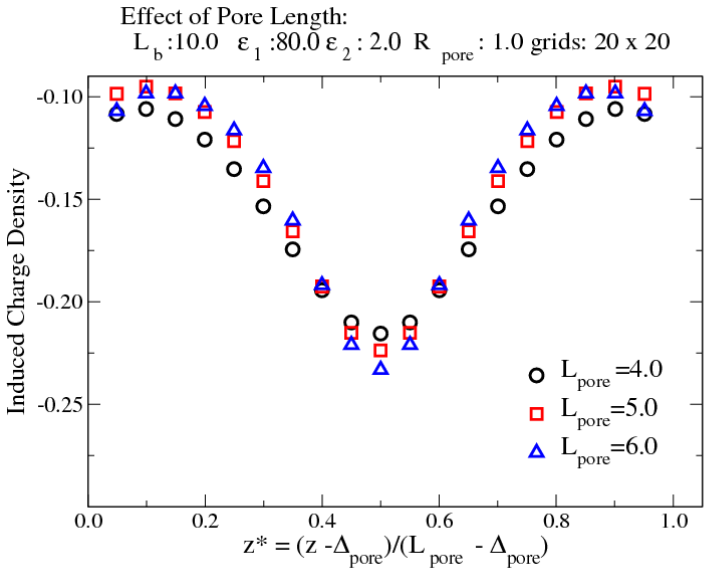


Figure 5.8: Polarization charge profile along the finite cylindrical channel with varying channel lengths, averaged over all azimuthal angles. The boundary element location along the channel is scaled (z^*) to be able to compare different channel lengths. The z values are shifted by $\Delta_{pore}/2.0$ before scaling, because of the restriction on scaling, $z \geq \Delta_{pore}$, while the initial boundary element is located at $\Delta_{pore}/2.0$

The surface polarization charge distribution induced by the source charge is averaged over all possible directions in the given side of the channel: along the cylindrical ($z-$axis) and radially ($r-$axis) along the membrane side. Fixed parameters were: $L_b = 10.0$, $\epsilon_1 = 80.0$, $\epsilon_2 = 2.0$, and $M = 20$, implying total of 400 surface charges on the cylindrical pore and on each membrane side.

In order to compare surface charge distributions for different values of L_{pore} and R_{pore}, the following scaling appears to be essential:

$$z^* = (z - \Delta_{pore})/(L_{pore} - \Delta_{pore}), z \geq \Delta_{pore},$$

$$r^* = (r - \Delta_{membrane})/(L_b/2 - \Delta_{membrane}), r \geq \Delta_{membrane}.$$

Results of ICCP^{3}M computations are presented in Figure 5.8. Here the induced charge density is plotted versus the scaled distance z^* for different values of pore length, namely $L_{pore} = 4.0, 5.0, 6.0$. One can see that the change in the pore lenght in the order of 20 percent does not induce significant changes in the induced charge density. However, in the middle of the pore, the induced charge grows with the pore length. In Figure 5.9 we plot the polarization density $\sigma(r)$ as a function of the scaled radius r^* for different values of the pore radius $R_{pore} = 1.0, 2.0, 3.0, 5.0$. There are two main points one can extract from this figure. First of all, the induced charge density decreases rapidly with increasing distance from the pore centre. The second finding is that, the more narrow the pore is the higher is the induced charge density.

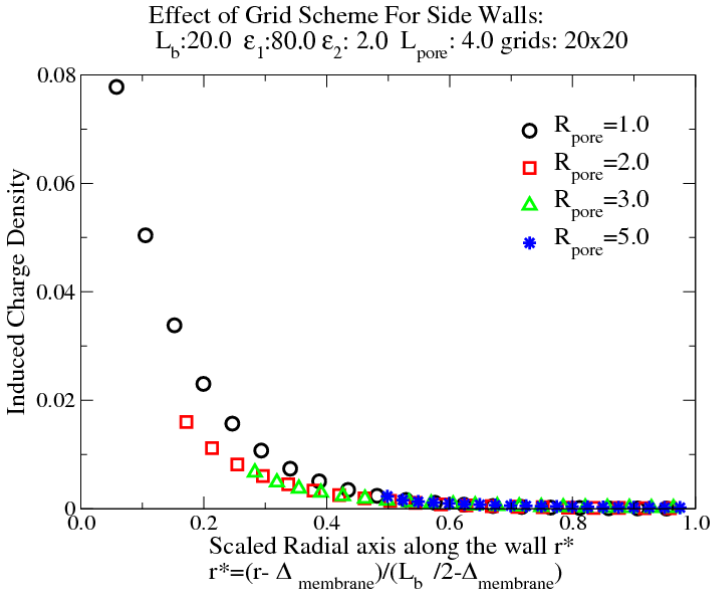


Figure 5.9: Polarization charge distribution along the radial membrane side versus the scaled distance from the center of the pore entrence to simulation box boundary is shown. The effect of varying pore radius is investigated.

5.4 Potential of Mean Force: Stiff DNA Model Inside a Nanopore

Polymer translocation in the context of macromolecular transport has been discussed in Section 5.1. However, most of the computer simulations presented there were focusing on a generic bead-spring model. No charges were assumed there, except for two models. In the first one, described in [113], the dielectric domains were approximated with an empirical scaling rather than obtained by solving the Poisson equation. In the other case, only the ejection of coions (ions that carry the charge same to DNA) was investigated [114].

On the contrary, in the present work we have successfully applied ICCP^{3}M to the translocation of a charged polymer, using the DNA model that we have developed and tested. It is quite crucial to understand how the total force acting on the polymer changes depending upon its position within the channel. This information provides an invaluable contribution into understanding translocation mechanisms.

The only analogous work known in the literature is ref [203], where the computation of the total force acting on the stiff, but unlike our case, uncharged polymer translocating through a narrow pore is studied in depth in the context of diffusion and ratchet mechanisms of translocation.

In our work, the DNA backbone charges were fixed along the cylindrical central pore axis. We have computed the effective force F_{CoM} on the center of mass of the stiff DNA. By fixing the DNA in different locations along the cylindrical central pore axis we extracted the *Potential of Mean Force* (PMF). The PMF profile is computed via the effective force F_{CoM} acting on the DNA fixed at different locations.

The very definition of the effective force F_{CoM} which acts on the given centre of mass located at

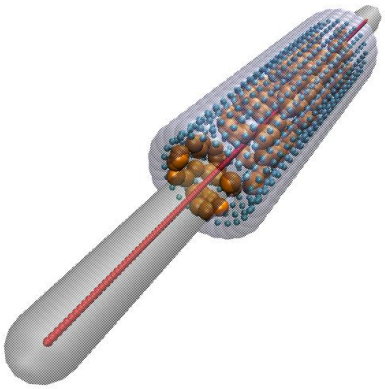


Figure 5.10: The rod DNA model in the channel geometry, when the DNA model is placed at the center of the channel. Areas depicted in gray represent the excluded volume interactions (van der Waals) against counterions, on the DNA rod and in the cylindrical channel. The charge centre of the DNA rod and the cylindrical channel are shown as red and blue respectively. The counterion initial conditions, are shown with corresponding van der Waals radii as orange.

z_{com} tells us the total force acting on the center of mass of the stiff rod. It is defined as a sum over the projection of the forces $\mathbf{f}_i$, where index i runs over the N rod charges, on the each backbone charge centre on the pore central axis $\hat{\mathbf{z}}$:

$$F_{CoM}(z_{com}) = \sum_{i=1}^{N} \mathbf{f}_i \cdot \hat{\mathbf{z}},$$

where $\mathbf{f_i}$ includes both electrostatic and short range contributions.

One can determine the PMF profile by measuring F_{CoM} at different z_{com} values, for example, starting from point a to point b, $a \leq z_{com} \leq b$, and integrate the resulting profile:

$$V_{pmf}(z_{com}) = \int_a^{z_{com}} dz'_{com} \left\langle F_{CoM}(z'_{com}) \right\rangle .$$

5.4.1 A stiff DNA force profile computed via Langevin Dynamics

The effective force and the PMF profile of the stiff DNA model developed in Section 5.2.3, are investigated via Langevin dynamics with and without induced charge computation.

The stiff DNA model is formed by 177 charges resulting in a 30 nm long rod. Channel and side walls are discretized with boundary elements, 2500 charges for the pore and 100 charges in a rectangular scheme on the side walls. We used a box length of 75 nm and pore length of 13.6 nm.

The initial configuration of the simulation, when the DNA is placed at the centre of the channel and its counterions are randomly distributed within the channel, is shown in Figure 5.10, where the centre of mass (CoM) of the DNA rod corresponds to 35.0 nm. Other CoM fixed locations vary from 17.5 to 59.5 (in box units). After fixing the CoM of a DNA rod at a given location in this interval, the effective force is measured. The sampling of the effective force is carried out after counterions and force values relaxed. Simulation parameters for the Langevin part are $T = 1.0$,

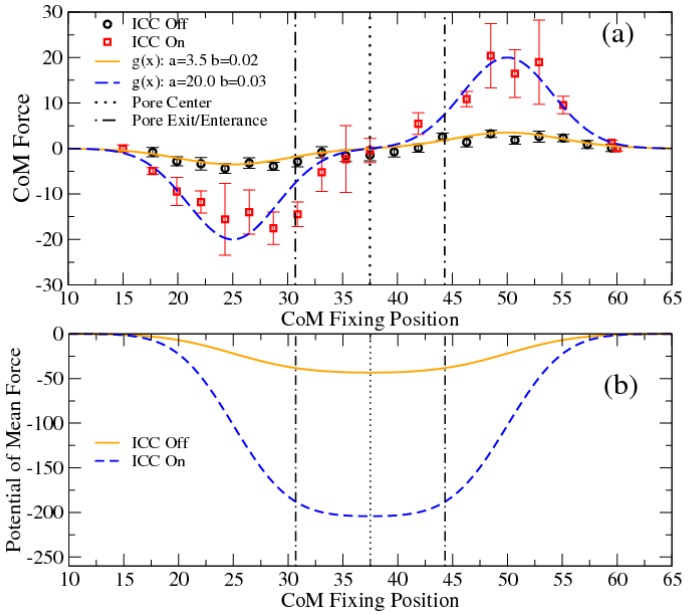


Figure 5.11: (a) The effective force profile of DNA rod translocation, with and without ICCP^{3}M . An emprical curve $g(z) = -a\exp\left[-b(z-25.0)^2\right] + a\exp\left[-b(z-50.0)^2\right]$, that describes the data quite well, is used. Coefficients (a,b) were employed: (20.0,0.03) and (3.5 and 0.02) for the simulation data with and without ICCP^{3}M respectively. (b) Corresponding Potential of Mean Force profile of DNA rod translocation for updates at every 3000 ICCP^{3}M , along a finite cylindrical pore with and without ICCP^{3}M (for updates at every 3000). Channel Centre is located at 37.5. Channel enterance and exit are located at 30.7 and 44.3.

and $\gamma = 0.5$. The procedure is repeated with the inclusion of dielectrics, where dielectric constants inside the channel and in the bulk were taken as 80 to mimic water, and 2 in the membrane sides, with a Bjerrum length of 0.71 nm. The ICCP^{3}M optimization procedure, discussed in Chapter 4, is updated every 3000 and 30000 steps. This will allow us to investigate the effect of updating frequency on the values of induced charges. Hence, three different profiles are obtained, with and without inclusion of dielectrics for two different updating frequencies, as shown in Figure 5.11 (a). The emprical curve

$$g(z) = -a\exp\left[-b(z-25.0)^2\right] + a\exp\left[-b(z-50.0)^2\right],$$

that describes the data quite well is used. Coefficients (a,b) were employed: (20.0,0.03) and (3.5 and 0.02) for the simulation data with and without ICCP^{3}M respectively. In all cases the effective force profiles are symmetric with rescpect to the channel centre due to inherent symmetry of the system under study. Extremum values in the profiles physically correspond to force barriers that the stiff DNA must overcome to be able to enter to the channel. The presence of surface polarization charges generates significantly higher effective force barriers at the entrance/exit of the pore (almost an order of magnitute higher), than in the absence of the induced charges, regardless of the ICCP^{3}M optimization procedure update. The location of the barrier is at the place where approximately two thirds of the DNA rod are inside the channel. This signifies the high electrostatic cost of *capture events*, during the diffusion of a stiff rod to the channel from bulk. The effective force approaches zero when DNA is completely removed from the channel as seen in Figure 5.11

(a).

The sign of effective force barrier heights shows a direction that is away from channel. This is due to the *dielectric repulsion* in the case of inclusion of dielectrics. In the case of no dielectric inclusion, barrier heights are smaller and appear only due to entropic contributions of counterions, confined within the channel.

The integral of the effective force profile lead to the PMF as defined in the previous section. It is shown in Figure 5.11 (b). The PMF profile gives us a clear idea about the height of the electrostatic barrier that DNA must overcome to translocate completely through the channel.

Hence, we demonstrate the striking effect of the Dielectric Boundary Force (DBF) during stiff DNA translocation and capture. Moreover, the importance of electrostatic bariers quantitatively demonstrated. Our results strongly support the notion that the DBF can not be ignored in any realistic modelling of polymer translocation.

Furthermore, a more refined way of modelling might also be useful, in the future to understand the underlying physics of ion channel selectivity and gating mechanisms in molecular simulations generally.

5.5 Summary

We have reviewed the importance of macromolecular transport and related recent research work that appeared in the literature from a coarse grained simulations point of view, that reflects our approach to DNA translocation. Coarse graining of DNA, membrane and solvent degrees of freedom were also outlined. While continuum electrostatics is taken as a basis for coarse graining of membrane and solvent, two different spatio-temporal models of DNA are described and compared. Specifically, force extension curves of analytical results from the WLC model and the ones obtained by our simulations in Chapter 3 are compared. These DNA representations are frequently used for analytical and computer modeling purposes. However, their shortcomings underlined the need for the development of a new coarse grained DNA model.

The suitability of our DNA model was tested by comparing counter-ion condensation around a stiff DNA within the cell model Poisson-Boltzmann (PB) calculations and Langevin dynamics simulations. Our results proved the introduction of overlapping excluded volume interactions on the backbone do not introduce any artifacts into the electrostatic characteristics of DNA.

The effect of different geometric parameters for the finite cylindrical channel and its sidewalls (membrane sides) for the computation of surface polarization charges due to a single source charge were discussed. The effect of pore length has been investigated with fixed simulation box length. We have observed that the smaller the ratio is, the smaller the induced charge is. Increasing the pore radius has resulted in smaller polarization, as expected, since the distance of the source charge to the dielectric interface increases.

Langevin dynamics simulations of translocation of stiff DNA through narrow pores have been employed for understanding the force and potential barriers, specially the one induced by the presence of a dielectric interface. Our quantitative measurements of the force profiles have shown that DNA must overcome a significant force barrier to place even one third of its length into the channel. This barrier completely disappears when DNA is completely out of the channel. Moreover, computing the potential of mean force enables us to quantify a potential barrier needed to be overcome by DNA to completely translocate from the channel centre. Once again, the striking effect of dielectric interface has been shown.

6 Conclusions and Outlook

This thesis was aimed at understanding the physical fundamentals governing Single Molecule Experiment (SME) techniques from two different perspectives.

We investigate the thermodynamics of stretching of biomolecules, such as DNA, and showed that the outcomes of force-spectroscopy measurements depend on measurement procedures (ensembles). It has been clarified and shown that ensemble equivalence is not possible even in the thermodynamic limit for a single chain stretching experiment with Langevin Dynamics simulations in the low force regimes. In contrast to the common understanding that ensemble equivalence prevails in the thermodynamic limit regardless of the force regime, the fact that the ensemble difference versus polymer length scaling does not hold is proved in this thesis.

A fundamental questions appeared when using SME techniques in the presence of confined geometries, such as the pulling of a DNA out of a nanopore: treating the membrane and solvent degrees of freedom accurately, **fast** and with sufficient spatio-temporal resolutions in computer simulations using continuum electrostatics. Hence, the mathematical and algorithmic problem of solving the Poisson equation in the case of inhomogeneous dielectrics is addressed. The existing solutions for this problem are reviewed in depth and the ICC algorithm is implemented and tested. A new, accurate and **fast** solution ICCP^{3}M , is proposed, implemented, compared with ICC and tested. The equivalence of ICC and ICCP^{3}M is shown analytically. Furthermore, the ICCP^{3}M is applied to understand the process of stiff DNA translocation.

Introducing Single Molecule Experiments and Biopolymers

The SMEs and basics of biopolymers from a coarse-grained computational soft matter and biological physics points of view have been reviewed and explored. The development of polymer physics, the fundamental characteristics and classifications of polymers, and major stretching SME techniques are outlined.

Question of Ensemble Equivalence

The construction of different possible measurement protocols in the family of SMEs, for example in force-microscopy, gives rise to the question of equivalence of different ensembles in these systems. In this context a study has been conducted of the equivalence of isotensional and isometric ensembles for single linear model polymers via coarse-grained Langevin Dynamics simulations. The notion that two ensembles are not equivalent in the zero and weak stretching forces for a Gaussian chain with zero and non-zero equilibrium bond lengths, even in the thermodynamic limiting case of very long chains has been shown, confirmed and enhanced. This conclusion has been inferred from quantitative force-extension data obtained via simulations. The force-extension curves for weak to intermediate regimes showed a *universal* behavior, when scaled accordingly, independent of polymer length. Moreover, we obtain the scaling of ensemble difference that are as a function of chain length for a wide range of force regimes. This study complements the the-

oretical understanding of fundamentals of thermodynamics of small polymeric systems as well as those of force-microscopy experiments.

Induced Charge Computation

Coarse-grained computer simulation of transport of macromolecules or ions through nanoscopic membrane channels are employed to understand the fundamental physics of DNA translocation. One of the physical properties inherent to membrane channels is their inhomogeneous electrostatic composition. The formulation of the Poisson problem in these systems has been shown generically by introducing polarization charge density on the dielectric boundary. In particular, matrix formulation of the problem and the Induced Charge Computation (ICC), are outlined, implemented and tested with simple analytically tractable problems. Moreover, the equivalence of the matrix formulation and the boundary element approaches were shown analytically. The derivation of an iterative formulation by using the Successive Over-relaxation (SOR) technique is also shown and justified physically.

The development of a novel iterative algorithm, ICCP^{3}M , that treats continuum electrostatic force computation in the presence of inhomogeneous dielectric domains accurately and fast, has been outlined, formulated, implemented and tested with simple analytically tractable geometries and compared to the Induced Charge Computation algorithm.

DNA Model

A coarse-grained model of DNA is needed for the simulation of stiff DNA translocation. For this reason, coarse-grained DNA models that appeared in the literature including those that are used for thermal denaturation, stretching and counterion condensation studies have been reviewed. In the lines of models used in counterion condensation studies, a coarse-grained DNA model has been developed further. This model is tested against cell model Poisson-Boltzmann calculations in stiff configuration, thereby concluding that the model does not generate any artifacts from an electrostatics point of view.

Stiff DNA Translocation Through Nanopores with Dielectrics

The combination of the DNA model and a developed dielectric algorithm enables us to perform simulations on stiff DNA translocation through nanoscale pores with the inclusion of dielectric interfaces. The potential of mean force profile of the stiff chain with or without dielectrics was computed by defining the force on the center of mass by means of coarse-grained Langevin dynamics simulations.

The results of these computations have shown the striking effect of dielectric boundary forces:

- The stiff DNA needs to overcome a significant force barrier to enter the pore. The height of this barrier is almost driven purely by dielectric interfaces.
- To completely translocate through the pore the stiff DNA needs to overcome another potential barrier at the center of the pore, the height of which is similarly determined by dielectric interfaces.

Further Outlook

The improvements or applications of themes that have been studied here are open for further development. The following itemized list shortly summarizes possible extensions:

- Extending the investigation of ensemble inequivalence to confined geometries. Possible comparison with theoretical works [204, 205].
- Extending the investigation of ensemble inequivalence to charged chains. There are pulling simulations of polyelectrolytes [206, 207, 208, 209] and a related experiment on condensation in DNA pulling [82].
- The construction of refined and improved WLC model of DNA with correct charge characteristics, follow the recent works [197, 198].
- Extending the ICCP^{3}M implementation to multiple and distinct closed dielectric domains. One possible route to solve this might be a self-consistent procedure, that repeats ICCP^{3}M routine separately for each different interface. One recent study in this direction is the used set of boundary integral equations [210].
- Measuring the DNA's effective charge via direct force measurements when the DNA is partly inside a nanopore with coarse-grained simulations including dielectric effects, and comparing them with experimental results [32].
- The exploration of stiff DNA translocation taking into account the effect of added salt, chain-length and curvature of the pore.
- Voltage assisted translocation of flexible DNA chain investigation and the measuring of the translocation times versus external field strength and chain length that are an intense debate in the literature [104, 103, 107, 108, 5, 109, 113, 109, 114, 189, 190, 191, 192, 111]. The effect of dielectric interfaces on the scaling arguments in translocation can be investigated.

7 Zusammenfassung

Die im letzten Jahrenzehnt erreichten Fortschritte in der Biologie, Biochemie und Physik haben wesentlich dazu beigetragen, dass es zu größeren Überschneidungen zwischen diesen Bereichen gekommen ist, von denen alle profitieren und die sich zum Teil synergetisch kombinieren. Eines der besten Beispiele für das Zusammenwirken neuer Techniken und Methoden sind Einzelmolekülexperimente (im Englischen "Single Molecule Experiment" (SME)) [1, 2, 3]. Basierend auf neuen experimentellen Methoden wie im Fall der Optischen Pinzette [70, 71, 1, 72, 73, 74, 75], Magnetischen Pinzette [80, 81, 82] oder der Atom-Kraft-Mikroskopie [65, 76, 77, 78, 8], und erweiterten theoretischen Behandlungen dieser Untersuchungsmethoden wurden diese Anwendungen ein zentraler Punkt der Forschung, die sich mit der Untersuchung der mechanischen Eigenschaften einzelner Moleküle beschäftigt und das oft sogar *in vivo* [10]. Das Verständnis der mechanischen Eigenschaften einzelner Moleküle erlaubt zum Beispiel die Bestimmung der freien Energie der Bindung molekularer Untereinheiten, wie zum Beispiel im Falle der erzwungenen Entfaltung von Proteinen [211, 212]. Das Interesse an der Manipulationen einzelner Molekülen beschränkt sich nicht nur auf ihre Eigenschaften in einer Umgebung, die abgesehen von der Messapparatur ziemlich leer ist. Im Vordergrund steht vielmehr die Untersuchung der Eigenschaften von linearen Polymeren, die mit Objekten wie synthetischen oder biologischen Nanoporen wechselwirken. Diese sind bei der molekularen Translokation in biochemischen Prozessen [180, 181, 104, 103, 5, 113, 109, 114] von Bedeutung, und sie haben auch ein Potenzial für eine neuartige Methode, die eine molekulare Erkennung [6] liefern könnte.

Die Kenntnis der Subtilitäten der statistischen Mechanik von einzelnen Molekülen ist daher die wichtigste Voraussetzung für den richtigen statistischen Ansatz zur Lösung von Problemen im Zusammenhang mit SME. Daraus ergibt sich das erste Ziel dieser Dissertation. Dieses liegt darin, die Equivalenz oder Inäquivalenz von Ensemblen im thermodynamischen Limes in SME zu untersuchen. Dieser Punkt ist für ein endliches System wie ein SME enorm wichtig. Zu unserer Überaschung ist er jedoch bisher in der statistischen Mechanik einzelner Moleküle noch nicht vollständig geklärt. Um weiterhin einen rigorosen Ansatz für die Simulation von SME in eingeschränkten Geometrien zu erhalten, ist es unerläßlich, dass eines der größten Defizite derartiger Simulationstechniken beseitigt wird, nämlich die korrekte Berechnung der Kräfte von geladenen Makromolekülen in der Nähe von dielektrischen Grenzen in vergröberten Polymermodellen zu gewährleisten. Deshalb widmet sich der zweite Teil dieser Arbeit der Entwicklung, Umsetzung und Erprobung eines effizienten und präzisen Algorithmus zur Berechnung dieser Kräfte, sogar im Falle von beliebig geformten Grenzen. Als Anwendung des Algorithmus, der im dritten Teil dieser Arbeit präsentiert wird, wurde das Potenzial der mittleren Kraft für ein starres DNA-Modell berechnet. Die Einzelheiten dieser drei Teile werden in den folgenden Abschnitten zusammengefaßt.

7.1 Ensemble–Inäquivalenz in Einzelmolekülexperimenten

Eine der wichtigsten Annahmen für SME besagt, dass im Falle von sehr langen Molekülen (wie Proteinen, DNA oder verschiedenen synthetischen oder biologischen Polymeren), die Ergebnisse der Messung nicht davon abhängen sollten, welche Kontrollparameter im Experiment angwandt werden [17]. Dabei entscheidet die Verwendung bestimmter Kontrollparameter, welches Ensemble im Experiment benutzt wird. Die Eigenschaft, dass der Erwartungswert für eine gemessene Observable in verschiedenen Ensembles zum gleichen Wert konvergiert, wird schließlich Ensemble–Äquivalenz genannt. Man kommt von der statistischen Mechanik zur Thermodynamik, indem man die thermodynamischen Potentiale durch die Zustandssumme des Systems definiert, wobei für alle extensiven Kontrollparameter der Grenzwert unendlich großer Werte gebildet wird. Im allgemeinen ist es dabei möglich, bei vorgegebenem statistischem Ensemble und Kontrollparameter ein konjugiertes Ensemble mit konjugiertem Kontrollparameter zu bilden, was über die Ableitung des thermodynamischen Potenzials des vorgegeben Ensembles geschieht. Dabei wird davon ausgegangen, dass im thermodynamischen Limes zwei konjugierte Ensembles gleichwertig sind. Das bedeutet, sie sollten die gleichen Erwartungswerte für die thermodynamische Größen produzieren [86].

Unter den vielen konjugierten Ensemblen, die sich in SME realisieren lassen, sind das isotensionale und das isometrische Ensemble wahrscheinlich am meisten verbreitet. Im isotensionalen Ensemble ist der Kontrolparameter die Kraft, die an einem Ende der Kette angelegt wird, und der Kontrollparameter des isometrischen Ensembles ist der Abstand zwischen den Endpunkten der Kette. Obwohl die Frage der Ensemble-Äquivalenz oft untersucht wurde, sowohl aus theoretischer Sicht [125, 126, 127, 128, 28, 29, 129, 130, 24, 131, 132, 17, 22, 19, 20, 133, 23], als auch unter Verwendung von Computer-Simulationen [134, 135, 136, 137, 138, 139, 140, 26, 21, 141], wurde bis jetzt keine schlüssige Sichtweise über die Ensemble–Inäquivalenz zwischen dem isotensionalen und dem isometrischen Ensemble für das Regime mit kleinen Kräften vorgestellt. Nach unserem Wissen existieren keine Arbeiten, die basierend auf Computersimulationen, zur Lösung dieser offenen Frage führen. In Experimenten, die einzelne Moleküle direkt manipulieren, wie im Falle der AFM oder der Optischen Pinzette, wirkt die Messapparatur auf den Terminus des Moleküls. Je nach Stärke der Wechselwirkung der Messappatur mit dem Molekül, wird ein Ende des Moleküls mit konstanter Kraft festgehalten oder die sich ergebende Kraft gemessen, während das andere Ende fest mit einer Oberfläche verankert ist [17]. Sowohl das isometrische als auch das isotensionale Ensemble können realisiert werden, indem die Kraftkonstante des Cantilevers des AFM oder die Federkonstante im Fall der Experimente mit der Optischen Pinzette [24] verändert wird, wobei die Anwendung sehr hoher Kräfte zu einer Realisierung des isometrischen Ensembles führt. Es bietet sich deshalb an, als Kontrollparameter den Vektor X zu nutzen, der durch die beiden Enden des gemessenen Objekts definiert wird, und seinen konjugierten Parameter, die Kraft F, die an einem Ende wirkt. Diese führt direkt zur Definition zweier entsprechender konjugierter Ensembles. In einem Fall fürt dies zum isotensionalen Ensemble, in dem eine konstante Kraft auf eines der Moleküls wirkt, und im anderen Fall zum isometrischen Ensemble, bei dem der Vektor X fixiert ist, wobei zu beachten gilt, dass wenn man nur den Abstand der Termini des Moleküls festhält, dies zu einem anderen Ensemble führen würde [28, 137, 138, 29, 30].

Ein mögliches Mittel, um die Äquivalenz der Ensemble in SME zu untersuchen, ist die Analyse der Kraft-Ausdehnungs-Kurven in den beiden Ensembles. Dies bedeutet, dass man sich die Graphen von $\langle F(X)\rangle$ im isometrischen Fall und $F(\langle X\rangle)$ im isotensionalen Ensemble anzuschauen sollte. Im Falle einer Ensemble–Äquvalenz sollten sich diese beiden Graphen im Limes unendlich

langer Ketten nicht unterscheiden. Neumann [30] hat gezeigt, dass bei einer korrekten Definition des Ensemble-Unterschieds Δ dieser in der Tat verschwindet, wenn die Zahl der Monomere gegen unendlich geht (zumindest für eine Gauß'sche-Kette) und man sich im Regime einer gemäßigten oder starken Dehnung der Kette befindet. Insbesondere sollte im Bereich der Überdehnung der Abstand der Molkülenden linear mit der Anzahl der Monomere skalieren, und der Ensemble Unterschied sollte sich gemäß $\Delta \approx (N-1)^{-1}$ verhalten [17]. Bei geringen Kräften zeigt sich jedoch ein deutlich anderes Verhalten der Ketten, und in diesem Bereich skaliert der Ensemble Unterschied nicht mit der Größe des Systems, sondern bleibt konstant. Für jede Kettenlänge, das heißt, für ein makroskopisches, wenn auch nicht unbedingt unendliches System, ist es dann möglich eine kleine Kraft zu finden, bei dessen Anwendung der Ensemble Unterschied nicht verschwindet und darüber hinaus auch nicht spürbar sinkt, selbst wenn die Kettenlänge vergrössert wird. Bei Ketten, die lang genug sind um die Gauß-Näherung zu erfüllen, kommt es dann zum einem Ergebnis, das unabhängig vom Modell der Kette ist.

In dieser Arbeit werden wir Ergebnisse von Simulation von Zugexperimenten an Polymeren vorstellen, die sich mit der Äquivalenz dieser Ensembles auseinandersetzen. Zudem soll die theoretische Einschätzung geprüft werden, dass im Falle einer schwachen oder verschwindend kleinen Kraft eine Ensemble–Äquivalenz im thermodynamischen Limes nicht möglich ist [30], da das übliche Skalierungsverhalten des Ensemble Unterschieds ($\propto 1/N$) im Vergleich zur Polymerlänge N in diesem Bereich nicht mehr gilt. Dieses Ergebnis steht im Widerspruch zu dem allgemeinen Verständnis, dass im thermodynamischen Limes, und unabhängig von der Kraft [130, 140, 17], immer Ensemble–Äquivalenz herrscht. Einige der Autoren in den oben zitierten Werken haben dadurch den Schluss gezogen, dass die Ensembles unendlich langer Ketten, beschrieben durch eine Gauß'sche Kette [130, 140], einer generische Kette [17] und Gitter Modellen [19, 20] äquivalent sind. Einige andere Autoren hingegen behaupten, dass es nicht zu einer Ensemble–Äquivalenz im thermodynamischen Limes für eine einzige Kette kommt [28, 137, 138, 29].

Zur Beantwortung dieser offenen Frage, der Äquivalenz von Ensemblen, wurden Computer-Simulationen zur Untersuchung des Verhaltens von linearen Polymeren in isometrischen und isotensionalen Ensembles durchgeführt, die im Prinzip die Resultate von Einzelmolekülexperimenten modellieren. Effekte, die sich aus der endlichen Größe des Systems ergeben, beeinflußen Δ in einer Größenordnung von $(N-1)^{-1}$, was für eine Gauß'sche Kette und starke Kräfte gezeigt wurde. Weiterhin wurde das Verhalten auch für Kugel-Feder-Modelle mit endlichem und verschwindendem Gleichgewichtsabstand überprüft. In einem Bereich, in dem geringe Kräfte angewandt werden zeigt sich ein dramatischer Wandel des Skalierungsverhaltens. Hier stellen wir fest, dass die Kraft-Ausdehnungs Kurven ein universelles Skalierungsverhalten aufweisen, das typisch ist für freie Gauß'sche Ketten im Gleichgewicht. Dies wiederum führt zu der Tatsache, dass im Falle verschwindend kleiner Kräfte nie eine Äquivalenz der Ensembles erreicht werden kann. Unsere Computer-Simulationen bestätigen und erweitern die Analyse von Neumann. Es wird aufgezeigt, dass Vorsicht geboten ist bei der Betrachtung der Thermodynamik einzelner Moleküle, die viele subtile Unterschiede im Vergleich zu Systemen im Volumen aufzeigt, trotz der formalen Analogien zwischen den beiden Systemen. Darüber hinaus hat sich gezeigt, dass der Übergang zwischen den Bereichen, in denen Inäquivalenz beziehungsweise Äquivalenz herrscht, ein erstaunlich breites Spektrum der Kräfte umfaßt. Hierzu gehört, dass die Skalierung sich bei weitem noch nicht wie $(N-1)^{-1}$ verhält, selbst bei reduzierten Kräften die nahe bei eins sind. Die Ergebnisse zeigen, dass die Ensemble–Inäquivalenz tatsächlich eine wichtige Rolle spielen kann und zwar nicht nur aus prinzipiellen Überlegungen, sondern auch aus der operativen Sicht der Messungen in Einzelmolekülexperimenten.

7.2 Die Berechnung der induzierten Ladung in begrenzten Geometrien

Ein fruchtbarer Ansatz zur Simulation der weich kondensierten Materie basiert auf der Reduzierung der Freiheitsgrade durch die Entfernung der uninteressanten Freiheitsgrade eines Systems. Dieses Verfahren wird als "coarse-graining" (Vergröberung) bezeichnet, und kann auf verschiedenen Ebenen durchgeführt werden. Eine der am häufigsten angewandten Methoden ist die Verwendung eines impliziten Lösungsmittels, welches durch die Anwendung einer Langevin-Gleichung auf die Polymerteilchen realisiert wird, die sowohl zufälligen Kräfte wie auch einen Reibungsterm enthält, der die Brownsche Bewegung der Lösungsmittelteilchen abbildet. Dadurch treten keine expliziten Lösungsmittel-Moleküle in der Simulation auf, was zu einer signifikanten Reduktion des Rechenaufwandes führt. Der Verlust des entropischen Beitrages zur freien Energie, der von den aus den Freiheitsgraden des Lösungsmittels hervorgeht, erfordert daher, dass die durch das Lösungsmittel vermittelte Wechselwirkung, wie zum Beispiel der hydrophobe Effekt, durch den Einsatz maßgeschneiderter effektiver Potentiale in das Modell eingearbeitet werden müssen. Im Rahmen des einfachsten dielektrischen Modelles, bei dem die dielektrischen Permittivität als homogen angenommen wird, gehen nicht nur die entropischen Effekte verloren, sondern es besteht auch keine Möglichkeit eine polarisierende Wirkung des Lösungsmittels in der Nähe von dielektrischen Grenzflächen in Betracht zu ziehen. Es wurden daher einige Verfahren entwickelt, um diese Lücke zu füllen, indem man die elektrostatische Wechselwirkung durch die numerische Lösung der Poisson Gleichung [45] erhält. Im Falle einer Simulation, in der eine dielektrische Grenzfläche vorhanden ist, kann man den Ansätz wählen, die Lösung der Poisson-Gleichung für ein äquivalentes Randwertproblem zu erhalten. Hierbei werden die Randbedingungen durch die Einführung von virtuellen Oberflächenladungen [48, 51, 55, 56, 49] realisiert, die von der Polarisation herrühren. Die Bestimmung der durch das elektrische Feld induzierten Oberflächenladungsdichte ermöglicht es somit, dielektrische Polarisationseffekte in Betracht zu ziehen. Die überwiegende Mehrheit der Algorithmen, die diesem Ansatz folgen, sind jedoch sehr rechenaufwendig, da sie mit dem Quadrat der Anzahl der diskretisierten Oberflächenelemente skalieren. Darüber hinaus ist es in der Regel nicht möglich, die Algorithmen auf periodische Randbedingungen anzuwenden. Wir präsentieren hier eine neue Hybrid-Methode (ICCP^{3}M), die das Problem der Berechnung der Kräfte an der dielektrischen Grenzfläche in einer Weise löst, die effizient, genau, und genügend schnell ist, um sie in Molekulardynamik-Simulationen zu verwenden. Der ICCP^{3}M -Algorithmus berücksichtigt exakt die periodischen Randbedingungen für beliebig geformte dielektrische Ränder. Der Algorithmus skaliert gemäß $\mathcal{O}(MlnM)$, wobei M die Summe aus der Anzahl der freien Ladungen und der Oberflächenladungselemente des Systems ist. Die Methode basiert auf einem iterativen Ansatz und stützt sich auf die Tatsache, dass sich die Positionen der Ladungen in einer Molekulardynamik-Simulation in der Regel bei einem Integrationsschritt nur geringfügig ändern. Daher kann die letzte Lösung für die induzierte Ladungsverteilung als Ausgangspunkt für die induzierte Ladungsverteilung im nächsten Integrationsschritt genutzt werden, was zu einer drastischen Reduzierung der Anzahl der Iterationen führt, die notwendig sind, um die Lösung bis zu einer vorgegebenen Genauigkeit zu konvergieren. Der Algorithmus wurde mit drei verschiedenen Geometrien erprobt, nämlich einer planaren, zylindrischen und kugelförmigen Grenzfläche. Da eine genaue Lösung (entweder analytisch oder numerisch) für diese Geometrien erhältlich ist, war es möglich, quantitativ die Genauigkeit der Methode zu überprüfen. Unsere Ergebnisse waren dabei immer in sehr gutem Einvernehmen mit der exakten Lösung. Abweichungen ergaben sich nur für die Fälle, dass das Testteilchen in einem so kleinen Abstand zur Oberfläche lokalisiert war,

der kleiner als oder vergleichbar mit der Gittergröße war, die zur Diskretisierung der Oberfläche genutzt wurde. Darüber hinaus zeigte sich, dass die ICCP^{3}M Methode, verglichen mit einer der alternativen ICC [58] Methoden, die vor kurzem vorgeschlagen wurde, nicht nur eine günstigere Skalierung aufweist, sondern auch im allgemeinen eine schnellere und genauere Performance im Falle der von uns getesteten Systeme aufweist.

7.3 DNA Translokation durch eine Pore: Dielektrische Effekte

Ein interessantes Problem, bei dem ein dielektrisches Medium als Grenzschicht vorhanden ist und das effektiv mit "coarse-grained"-Strategien angegangen werden kann findet man bei Betrachtung der DNA-Translokation durch eine Nanopore. Eines der einfachsten, aber dennoch nicht-trivialen, vergröberten Modelle für eine Doppelhelix-DNA besteht aus einem steifen, geladenen Stab, um den ein bestimmtes Volumen für alle Teilchen definiert wird, und der von seinen Gegenionen umgeben ist. In unseren Simulationen wurde die DNA stets geradlinig gehalten, jedoch kann man die Steifigkeit im Prinzip über ein Potential variieren, welches an die Kopplung der Basenpaare angreift, die durch eine Punktladung modelliert werden. Die Regionen des ausgeschlossen Volumens der Basenpaare überlappen beträchtlich. Diese Überschneidungen werden genutzt, um die Oberfläche der DNA, die für die Gegenionen zugänglich ist, zu modellieren und damit die geometrischen Eigenschaften der doppelsträngigen DNA festzulegen.

Die Verwendung eines derartig einfachen Modells ist gerechtfertigt, so lange die DNA-Fragmente kürzer als 50 nm sind, da dies der Wert der Persistenzlänge unter physiologischen Bedingungen ist, also der Bereich, auf dem die DNA sich wie ein Stab verhält. Außerdem braucht man eine explizite Darstellung der Gegenionen, um zu berücksichtigen, dass Korrelationen zwischen den Ladungen im Prinzip auftreten können. Dies kann man sogar im Fall von monovalenten Gegenionen erwarten, da eine begrenzende Umgebung vorhanden ist. Dieses Modell kann daher einen realistischen Einblick in die Rolle der Gegenionen bei der Wechselwirkung zwischen Pore und DNA an der dielektrischen Grenze liefern. Die anderen Parameter des DNA-Modelles wurden in einer Weise gewählt, um neben den geometrischen auch die elektrostatischen Eigenschaften einer echten Doppel-Helix DNA darzustellen. Ein erster Test wurde durch einen Vergleich der Ergebnisse unseres Modells mit einwertigen und zweiwertigen Gegenionen und einem Zellenmodell [202], dass in der Poisson-Boltzmann-Näherung berechnet wurde, für verschiedene Werte des Manning-Parameters (von z = 1.0 bis z = 4.2) durchgeführt. Wie erwartet, wies unser Modell eine sehr gute Übereinstimmung mit der Poisson-Boltzmann-Vorhersage für den Fall monovalenter Gegenionen und dem kleinsten Manning-Parameter auf. Dieser Test hat uns klar gezeigt, dass unser Modell keine Artefakte mit dem ausgeschlossenen Volumen in die vorhergesagten elektrostatischen Eigenschaften der DNA einführt.

Als nächstes haben wir das Potentials der mittleren Kraft (PMF) in einer dielektrischen Nanopore für unser DNA–Modell berechnet. Es sollte nicht unerwähnt bleiben, dass viele der "coarse-grained" Simulationsstudien, die sich mit makromolekularem Transport beschäftigen, die dielektrischen Randbedingungen nicht korrekt behandeln. Einige der Modelle vernachlässigen entweder die dielektrische Diskontiuität komplett, behandeln diese mit einer semi-empirischen Skalierung [113] oder verwenden in der Simulation eine ungeladenen Kette. Nur in einem einzigen Fall wird die dielektrische Diskontinuität genau berücksichtigt [114], jedoch liegt der Fokus der Studie auf den Eigenschaften der Gegenionen die aus der Pore abgestoßen werden, wenn das DNA-Molekül in der Mitte der Pore festgehalten wird. In unserem Fall wurden der Einfluß der dielek-

trischen Diskontinuität, die von der Porenoberfläche ausgeht, mit dem ICCP^{3}M -Algorithmus korrekt berücksichtig. Als Reaktionskoordinate zur Berechnung des PMF des DNA-Fragments wurde dabei dessen Schwerpunkt benutzt. Tatsächlich zeigte sich, daß die Wechselwirkung zwischen Pore und DNA, die auch durch die Gegenionen modifiziert wird, erheblich von Polarisationseffekten abhängt, die durch die dielektrische Diskontinuität induziert werden. Tatsächlich ist die Höhe der Barriere der freien Energie am Eingang der Pore deutlich erhöht, wenn die Wirkung der induzierten Oberflächenladungen mit in Betracht gezogen wird. Diese erste Studie zeigte eine deutliche Wirkung der induzierten Kräfte an der dielektrischen Grenze. Deshalb ist dieser Effekt zu berücksichtigen, wenn mit Hilfe von "coarse–grained" Modellen die Physik der DNA-Translokation, Ionenkanal-Selektivität oder allgemein Signalübertragung durch geladenen Massentransport durch Nanoporen untersucht wird.

Zusammengefaßt setzt sich diese Arbeit mit wichtige Themen der Physik von einzelnen Molekülen auseinander. Dabei reicht der Inhalt von der grundlegenden Frage der Ensemble–Äquivalenz in der Statistischen Mechanik von Einzelmolekülexperimenten bis hin zur Entwicklung einer allgemeinen Methode zur Behandlung dielektrischer Diskontinuitäten in periodischen Geometrien, i.e. des ICCP^{3}M Algorithmus. Dieser Fortschritt im Bereich der Methoden erlaubte es uns, das Potential der mittleren Kraft eines Modells eines DNA-Segments in einer Nanopore zu bestimmen. Somit ebnet dieser Erfolg den Weg für eine neue Generation von vergröberten Simulationen an geladenen Systemen der weichen Materie, die sich in der Nähe von dielektrischen Grenzflächen befinden.

A Summary of Implementations

Here, the general summary of implementation of Induced Charge Computation (ICC) and ICCP^{3}M in ESPResSo package [61] is sketched. The central idea of P^{3}M is summarized in Section A.1. The main code parts of ICC and ICCP^{3}M are outlined in Sections A.4 and A.2, respectively. Some of the planar and cylindirical dielectrics ESPResSo scripts developed in testing ICCP^{3}M implementation are presented in Section A.3. Enhanced comments in codes is placed as much as possible.

A.1 The Central Idea of P^{3}M

Solving $N-body$ problem is one of the central theme in computational physics. One of the novel method to address this problems is particle-particle particle-mesh (P^{3}M) algorithm. There are excellent references that studies the details of this algorithm [40, 60]. Here we only outline the central idea of splitting electrostatic interactions with long and short range parts. So we compute the short range directly and long range part in the discrete Fourier space.

1. Assign charges into a k-space mesh (Discretization error).
2. Solve Poisson Equation on the mesh (Aliasing problem).
3. Calculate lattice forces from energies (Discretization error in derivative and 3 FFTs).
4. Back interpolation of forces to particles (Interpolation errors).

P^{3}M has a complexity of $\mathscr{O}(NlogN)$ instead of $\mathscr{O}(N^{5/3})$ of Ewald sum method.

Code Box 1 ***The C structure of the ICCP^{3}M module in ESPResSo***

```
/* iccp3m data structures*/
typedef struct {
    int last_ind_id;   /* Last induced id can not be smaller then 2 */
    int num_iteration; /* Number of max iterations                  */
    double e1;         /* Dielectric constants                      */
    double e2;
    double area;        /* Area of the grid element */
    double *areas;      /* Array of area of the grid elements */
    double convergence; /* Convergence criterion */
    double *nvectorx,*nvectory,*nvectorz; /* Surface normal vectors */
    int selection;  /* by default it is not selected*/
    double relax;   /* relaxation parameter for iterative */
    int update;         /* iccp3m update interval */
    double *fx,*fy,*fz; /* forces iccp3m will use*/
  } iccp3m_struct;
```

A.2 Serial Implementation within ESPResSo Package

The implementation designed in such a way that method can be used as general as possible. For example, moving dielectric boundary, multiple grid resolution, controlling maximum number of iteration, relaxation constant and convergence criterion of the procedure and update interval. Also, user must be able to extract the information that how many iteration steps have taken for convergence. The primary data structure of ICCP^{3}M is shown in Code Box 1.

Code Box 2 ***ICCP^{3}M command specification in ESPResSo .***

```
iccp3m <last_ind_id> <e1> <e2> <num_iteration> <convergence>  \
                <relaxation> <area> <normal_components> <update>
 <last_ind_id>  :  Image charges id's start from 0 to last_ind_id
 <e1>           :  Dielectric constant of the medium that source
                   charges reside
 <e2>           :  Dielectric constant of the other medium
                   (i.e. membrane)
 <num_iteration>:  Number of iterations
                   before routine gives up updating source charges
 <convergence>  :  Convergence criterion
 <relaxation>   :  Relaxation parameter.
                   It must be in between 0 and 2.
 <area>         :  TCL array of all boundary element
                   areas (a0,a1,a2,...)
 <normal_components>: TCL array of Components of normal
                   vectors for all boundary
                   elements (x0,y0,z0,x1,y1,z1,x2,y2,z2,...)
 <update>       : update time step (not implemented, reserved)
```

Specially, no single line of P3M code is touched in the original *p3m.c*. If needed a function from P^{3}M , instead it is copied and modified in ICCP^{3}M with a new name (with a suffix *_iccp3m*). This will allow future developers of P^{3}M and ICCP^{3}M not to interfere each other.

In this initial implementation correctness is taken as primary target and performance as a secondary. So there are two new files called *iccp3m.c* and *iccp3m.h*. Almost every piece of code needed to perform ICCP^{3}M is there, except the command call, which is in *initilize.c*. The ICCP^{3}M command, see Code Box 2, must be applied whenever induced charge values need an update. The current implementation allows a dielectric boundary to change its shape dynamically on the fly.

Note that we haven't given details of $force\_calc\_iccp3m()$ function where it goes in to P^{3}M implementation. The main iteration function is called $iccp3m\_iterate()$. Initial step in ICCP^{3}M iterations is determining electrostatic forces on the induced charges. This is achived by the function $force\_calc\_iccp3m()$, and source-source charge interaction has not taken into account while it isn't needed during iteration.

Code Box 3 *Core ICCP^{3}M Iteration Code in ESPResSo*

```
  for(j=0;j<iccp3m_cfg.num_iteration;j++) {
       force_calc_iccp3m(); /* Calculate electrostatic forces
                               excluding source source
                               interaction*/
      diff=0;
    for(c = 0; c < local_cells.n; c++) {
      cell = local_cells.cell[c];
      part = cell->part;
      np   = cell->n;
      for(i=0 ; i < np; i++) {
          if(part[i].p.identity <= iccp3m_cfg.last_ind_id) {
                /* interface grids */
               fdot=part[i].f.f[0]*iccp3m_cfg.nvectorx[part[i].p.identity]+
                   part[i].f.f[1]*iccp3m_cfg.nvectory[part[i].p.identity]+
   part[i].f.f[2]*iccp3m_cfg.nvectorz[part[i].p.identity];
                hold=part[i].p.q/iccp3m_cfg.areas[part[i].p.identity];
                qold=part[i].p.q;
hnew=hold;
              if(hold != 0) {
                 fdot=fdot/(2*3.14159265); /* divide fdot with 2pi */
                hnew=hold + iccp3m_cfg.relax*((del_eps*fdot)/qold-hold);
              }
                difftemp=fabs(hold-hnew);
                if(difftemp > diff) { diff=difftemp; } /* Take the largest
                                                          error for convergence */
                part[i].p.q=hnew*iccp3m_cfg.areas[part[i].p.identity];
                  if(fabs(part[i].p.q) > 100) {
                    char *errtxt = runtime_error(128 + 2*TCL_DOUBLE_SPACE);
                   ERROR_SPRINTF(errtxt, "{error occured 990 : too big charge
                                          assignment in iccp3m!
    q >100 , normal vectors or computed forces might be wrong or
                            too big! assigned charge= %f } \n"
    ,part[i].p.q);
                     break;
                   }
           }
       }  /* cell particles */
    } /* local cells */
      printf(" iccp3m iteration j= %d convergence_cre = %f \r",j,diff);
      if(diff < iccp3m_cfg.convergence) {
printf("ICCP3M converged step=%d \n",j);
break;
      }
  } /* iteration */
```

ESPResSo particle and cell structure is used in accessing to particle properties such as the charge and the force on it, see Code Box 3. Total forces on the particles were stored separately before ICCP^{3}M iteration and revived afterwards. Because, the procedure updates forces (purely electrostatics) on the induced charges, where they are real particle in ESPResSo .

A.3 ICCP^{3}M Test Scripts

In this section we present some of the ESPResSo scripts we developed in testing ICCP^{3}M .

A.3.1 Planar Dielectric Interface

The example is studied in dept in Section 4.6.1: A single charge sitting infront of a planar dielectric interface seperating two dielectric region. The situation is shown in the Figure 4.5a. The tcl script is as follows:

Code Box 4 *Planar Dielectrics: System parameters*

```
#
# Mimic Boda et al PRE 69, 046702 (2004)
#    Bjerrum  Length of 2.321
#
set name  "plane_dielectric"
set skin         0.2
set temp         1.0
set time_step    0.01
set pi           3.141592653589793
set accuracy     1.0e-5

# Other parameters
set tcl_precision 8
# Arguments that adjust
set args [llength $argv]

if {$args != 11} {
 puts "Something wrong with args; usage is
 Espresso plane_dielectrics.tcl charge_number_x charge_number_y e1 e2 lb\n"
 puts "distance_source file_name box_l positionX positionY \n"
 exit 0
}

set number_x [lindex $argv 1]  ;# number of induced charges on x-direction
set number_y [lindex $argv 2]  ;# number of induced charges on y-direction
set distance [lindex $argv 3]  ;# distance of the source charge from the plane
                               ;# (z-coordinate of the source charge)
set filename [lindex $argv 4]  ;# file
set box_l    [lindex $argv 5]  ;# Simulation Box Length
set positionX [lindex $argv 6] ;# x-coordinate of the source charge
set positionY [lindex $argv 7] ;# y-coordinate of the source charge
set e1 [lindex $argv 8] ;# dielectric constant of the first domain
set e2 [lindex $argv 9] ;# dielectric constant of the second domain
set lb [lindex $argv 10] ;# Bjerrum length
```

System parameters must be provided initially, see Code Box 4, such as source charge coordinates, the number of induced charges on the planar dielectric boundary and appropriate dielectric constants.

Code Box 5 ***Planar Dielectrics: Drived parameters***

```
set box_l [expr double($box_l)]
# Simulation Box
setmd box_l            $box_l $box_l $box_l
set center             [vecscale 0.5 [setmd box_l]]
setmd periodic         1 1 1
setmd time_step        $time_step
setmd skin             $skin
#
set filename "$filename$number_x$number_y"
set number_induce [expr 2*int($number_x*$number_y)]
set induce_q [expr -1.0/$number_induce]
set delta_x [expr $box_l/$number_x]
set delta_y [expr $box_l/$number_y]
set del_area [expr $delta_x*$delta_y]
puts "delx= $delta_x dely=$delta_y delarea=$del_area\n"
# Set up plane particles (induce grids)

# First dielectric plane
set z1 [expr $box_l/4]

# source particle
set distance [expr $distance+$z1]
puts "grid  $number_x by $number_y distance= $distance \n"

# Second dielectric plane
set z2 [expr $box_l-$box_l/4]
```

Drived simulation parameters are given, see Code Box 5,for example number of the boundary elements, total number of induced charges.

The second dielectric plane is introduced due to the necessity of preventing ill defined dielectric regions. Otherwise regions on the simulation box, that are in perpendicular direction to the planes, boundaries would not be continuous, that is to say dielectric region would change on the box boundary. Introducing the second dielectric plane that divides the dielectric region having dielectric constant ϵ_1 from the second region having dielectric constant ϵ_2 generates a continous dielectric region of having dielectric constant ϵ_2. This situation is depicted in Figure 4.5b.

Placement of induced charges for both planes carried out on a rectangular boundary element configuration, see Code Box 6. Every induced charge is located at the center of the boundary element. Hence, each charge on the dielectric planes represents an area. Assignment of area of boundary elements and corresponding unit vectors is carried out when a charge is placed with the $part$ command. While the boundary elements are regularly placed, area of each elements is the same. However direction of normal vectors are reversed for the second dielectric plane. Because, normal vectors should point to the region where source charges are located. Lists $areas$ and $normal$ contain the area and unit vectors sequencially for boundary elements that is compatable to $iccp3m$ command. Also, induced charge values can be assigned randomly, where the sum of the values should match with the source charges values, which is 1.0 here. So, we assigned a constant value $\$induce\_q$.

Following the placement of induced charges and their properties, source charge is placed in

between two dielectric planes where the region that has a dielectric contant of ϵ_1.

Code Box 6 *Planar Dielectrics: Induced Charges Placement*

```
# Decorate 1st plane
set x [expr $delta_x/2]
set k 0
for {set i 0} {$i < $number_x} {incr i} {
  set y [expr $delta_y/2]
  for {set j 0} {$j < $number_y} {incr j} {
    part $k pos $x $y $z1 fix 1 1 1 q $induce_q type 1
    set nx 0.0
    set ny 0.0
    set nz 1.0
    lappend normal $nx
    lappend normal $ny
    lappend normal $nz
    lappend areas $del_area
    set y [expr $y+$delta_y]
   # puts "Setting surface particle $k q=$induce_q"
    set k [expr $k+1]
  }
    set x [expr $x+$delta_x]
}
# Decorate 2nd plane
for {set i 0} {$i < $number_x} {incr i} {
  set y [expr $delta_y/2]
  for {set j 0} {$j < $number_y} {incr j} {
    part $k pos $x $y $z2 fix 1 1 1 q $induce_q type 1
    set nx 0.0
    set ny 0.0
    set nz -1.0
    lappend normal $nx
    lappend normal $ny
    lappend normal $nz
    lappend areas $del_area
    set y [expr $y+$delta_y]
   # puts "Setting surface particle $k q=$induce_q"
    set k [expr $k+1]
  }
    set x [expr $x+$delta_x]
}

set assigned_k $k
if {$number_induce != $assigned_k} {
 puts "Something wrong with assigned plane particle ids:
 number of induced charge $number_induce but assigned is $assigned_k\n"
 exit 0;
}
if {$distance < $z1 || $distance > $z2 } {
 puts "Distance for source can not be smaller then $z1 and
 greater then $z2 but  Distance = $distance\n"
 exit 0;
}
```

Code Box 7 ***Planar Dielectrics: Electrostatics and ICCP^{3}M***

```
# set source particle
set source_id $assigned_k
set sx [expr $box_l/2+$positionX]
set sy [expr $box_l/2+$positionY]
set sz [expr $distance+$z1]
    puts "Setting source particle $source_id"
part $source_id pos $sx $sy $distance q 1.0 type 1
writepdb config.pdb
# Langevin thermostat
thermostat off
# P3M tune
   # puts "p3m tune \n"
   # inter coulomb $bjerrum p3m tunev2 accuracy $accuracy r_cut 0.0 mesh 0 cao 0
   # set p3mparams [inter coulomb]
   # puts "p3m done: $p3mparams \n"
   # exit
# 25x25
#inter coulomb 1.3 p3m 3.5511093 16 4 0.57177944 9.9125964e-06
set accuracy 1e-4
inter coulomb $lb p3m tunev2 accuracy $accuracy r_cut 0.0 mesh 32
set p3mparams [inter coulomb]
puts "p3m done: $p3mparams \n"

set time [exec date +%s ]
puts "iccp3m go..time: $time and number
      induce= $number_induce del area=$del_area \n"
set induce_last [expr $number_induce-1]
#Usage: iccp3m <last_ind_id> <e1> <e2> <num_iteration> <convergence>  \\
#              <relaxation> <area> <normal_components> <update>
# iccp3m cascade
 for {set i 0} { $i < 5 } {incr i} {
  iccp3m $induce_last $e1 $e2 1000 1e-7 0.8 $areas $normal 1
  puts "iccp3m config done \n"
  integrate 0
  inter coulomb $lb p3m tunev2 accuracy $accuracy r_cut 0.0 mesh 32
 }

integrate 0
puts "integration finished   !"
# report force on the source particle at a distance
set sforce [part $source_id print f]
puts  "$distance $sforce"
set infile [open $filename a+]
puts $infile "$distance $sforce"
close $infile

exit
```

After initial P^{3}M tuning, ICCP^{3}M procedure is repeated in cascades of re-tuning of P^{3}M at each cascade and recomputing forces on each surface charges with $integrate$ command, see Code Box 7. This procedure yields to a stable induced charge values, see Section 4.6.4.

A.3.2 Cylindrical Dielectric Interface: Two Charges

An opposite charge pair inside cylindrical dielectric interface is studied in Section 4.6.4. Here we describe the ESPResSo script of this example.

Code Box 8 *Cylindrical Dielectric Interface: Reporting Functions*

```
#   ICCP3M for 2 charges +/- to get forces on them.
#   With varying Bjerrum Length, seperation distance,
#   box_length and pore radius.
# - Report polarization density as well
#
########################################## Procedures Start
#
# Polarization density along phi=0
#
proc print_polarization_density {induce_id  del_area  bjerrum  box_l  radius}  {
puts "    Now result is being written to density_iccp3m.dat"
set fdensity [open "density_iccp3m_$bjerrum.dat" "w"]
     set along [expr 0.5*$box_l+$radius]
     puts $fdensity "# Density Along phi=0"
     puts $fdensity "#   z-axis     Value_Charge_density "
   for {set i 0} {$i<=$induce_id} {incr i} {
     set pos [part $i print pos]
     set x [lindex $pos 0]
     set z [lindex $pos 2]
     if { $x == $along } {
         set value_q [part $i print q]
 set value_density [expr $value_q/$del_area]
         set z0 [expr $z]
 puts $fdensity "$z0  $value_density"
    }
  }
  close $fdensity
}

#  Determine Total Charge
#    Returns total charge in the whole system
proc systotalcharge {} {
   set tot [setmd n_part]
   set tcharge  0.0
  for {set i 0} {$i <$tot} {incr i} {
     set tempq [part $i print q]
     set tcharge  [expr $tcharge+$tempq]
   }
 return $tcharge
}
```

Extracting polarisation charge distribution along cylinder central axis is achieved by the procedure $print\_polarization\_density$. Initial charge neutrality of the system can be confirm with the procedure $systotalcharge$. These functions are given in Code Box 8.

Code Box 9 *Cylindrical Dielectric Interface: Placing Induced Charges*

```
# Place Induced Charges on a cylinder, find normal vectors and areas
#   Given Radius, Number of charges in r and z direction, box_length
#   (Notice that global variables normal and areas are defined in the proc)
proc place_induced {radius nr  nz   box_l  total_source}  {
    set pi  3.141592653589793
    set del_nr    [expr 2*$pi*$radius/$nr]
    set del_nz    [expr $box_l/$nz]
    set del_area  [expr $del_nr * $del_nz]
    set del_phi   [expr 2*$pi/$nr]
      if { $nz % 2 == 1 } { set nz "[expr $nz+1]" }
       set np [setmd n_part] ; set k 0 ; set temp 0.0; set qt 0.0;
       for {set i 1} { $i < [expr $nr+1] } {incr i} {
              set phi [expr $del_phi * $i]; set cosphi [expr cos($phi)] ;
              set sinphi [expr sin($phi)]
             for {set j 0} { $j <  $nz } {incr j} {
                 set qx [expr 0.5*$box_l+$radius*$cosphi]
                 set qy [expr 0.5*$box_l+$radius*$sinphi]
                 set qz [expr $j*$del_nz]
         set temp [expr rand()*0.1]  ;# random number between [0.0,0.1]
         # set particle on a cylindrical interface
         part $k pos [expr 0.5*$box_l+$radius*$cosphi]  \
                           [expr 0.5*$box_l+$radius*$sinphi]  \
                           [expr $j*$del_nz] q $temp type 1 fix 1 1 1
                  set qt [expr $qt+$temp]
          incr k
                }
           }
                    set induce_id [expr $k-1] ; set total 0.0;
                    set qavg [expr ($total_source-$qt)/double($k-1)]
                    for {set i 0} { $i < $induce_id} {incr i} {
                         set org_q [part $i print q]
                         set org_q [expr $org_q-abs($qavg)]
                         part $i q [expr -1.0*$org_q]
                         # part $i q 0.0
                         set total [expr $total+$org_q]
                       }
      # Generate Grid (Boundary Element) Unit Vectors and Areas
                 global normal ; global areas
                for {set i 1} { $i < [expr $nr+1] } {incr i} {
                  set phi [expr $del_phi * $i]
                  set cosphi [expr cos($phi)]
                  set sinphi [expr sin($phi)]
                   for {set j 0} { $j < $nz } {incr j} {
                         set x [expr $cosphi]
                         set y [expr $sinphi]
                         set z 0
                         lappend normal $x ; lappend normal $y ;
                         lappend normal $z ; lappend areas $del_area
                         incr k
                       }
                 }
      return $induce_id
}
```

Code Box 10 *Cylindrical Dielectric Interface: Simulation Parameters*

```
#        Force on two charges inside the center of dielectric cylinder
#  under different parameters i.e. bjerrum length, radius,
#  box_length, charge, seperation
      thermostat off ;  set pi  3.141592653589793;
      set accuracy 1e-5  ;# p3m
      set e1 80.0;  set e2 2.0;
      set iteration_giveup 1000   ; set convergence_accuracy 1e-8
      set relaxation_constant 0.6 ; set time_step 0.01
      set skin 0.2   ;
      set box_l 20.0 ;# initial box length to adjust initial radius
      set bjerrum 0.71
      set max_cycle 5
 # Now apply the given parameter ranges
      puts "Box Length = $box_l";
              setmd box_l  $box_l $box_l $box_l
              setmd periodic 1 1 1
              setmd time_step $time_step
              setmd skin $skin
                      set radius !.0
      set nr 8
      set nz 8
        set del_nr        [expr 2*$pi*$radius/$nr]
        set del_nz        [expr $box_l/$nz]
        set del_area      [expr $del_nr * $del_nz]
        puts "del_area = $del_area ";
        ;# Total value of the source charges, +/- charge
        set total_source 0.0

       ;# Place Induce Charges on the cylindrical interface
set last_induce [place_induced  $radius $nr $nz $box_l $total_source]
set id0 [expr $last_induce+1] ;# source particle ids
set id1 [expr $id0+1]
        set dist 5.0 ;# The seperation between charges will be 10.0
puts "Particle ids; $id0 and $id1, dist $dist";

       ;# The ratio beween the box length and radius of
        # the cylindrical interface
set RoverL [expr double($radius/$box_l)]
puts "RoverL= $RoverL";

      # Source Charges placement
      part $id0 pos [expr 0.5*$box_l] [expr 0.5*$box_l]  \
                    [expr 0.5*$box_l+$dist] q -1.0 type 0 fix 1 1 1
      part $id1 pos [expr 0.5*$box_l] [expr 0.5*$box_l] \
                    [expr 0.5*$box_l-$dist] q 1.0 type 0 fix 1 1 1

      writepdb config.pdb ;# write the system configuration.

       puts "bjerrum= $bjerrum RoverL=$RoverL  \
             RoverL/bjerrum=[expr double($RoverL/$bjerrum)]";
```

The placement of induced charges on the cylinder of given radius and grid scheme is achieved by the procedure $place\_induced$, see Code Box 9. This procedure assigns the normal vector and area of the each boundary element in a tcl list compatable to $iccp3m$ command. The rest of the simulation parameters and placement of opposite charges on the central cylinder axis is summarized in Code Box 10.

Code Box 11 *Cylindrical Dielectric Interface: Iteration Cascade*

```
# ICCP3M - P3M Optimization Cascades

for {set cycle 0} {$cycle < $max_cycle } {incr cycle} {
   puts "optimising p3m for ;  \
         number of charges =[expr $id1+1] "
   inter coulomb $bjerrum p3m tunev2 accuracy $accuracy r_cut 0 mesh 32 cao 0
   puts "optimising p3m done : [inter coulomb]"

   integrate 0 ;# Now Get the forces on all charges from P3M
   set iteration [iccp3m $last_induce $e1 $e2 $iteration_giveup \
                 $convergence_accuracy $relaxation_constant $areas $normal 1]

   puts "-- iccp3m converged in $iteration "
   puts "iccp3m exited without errors ! at Half the seperation $dist and  \
        Bjerrum Length $bjerrum  R/L=$RoverL "

   integrate 0 ;# Now Get the forces again
   set f [part $id1 print f]
   set fz [lindex [part $id1 print f] 2]
   set tot 0.0

  for {set ms 0} {$ms < [setmd n_part]} {incr ms} {
     set qtemp [part $ms print q]
     set tot [expr $tot+$qtemp]
    }
     set f [part $id0 print f]
     set fz [lindex [part $id0 print f] 2]
     print_polarization_density $last_induce  $del_area  \
      $bjerrum $box_l $radius

 } ;# cascade ends
     set fz00 [lindex [part $id0 print f] 2]
     set fz11 [lindex [part $id1 print f] 2]
     puts "Charge Seperation=[expr 2*$dist]\t  Fz0= $fz00 \t \
            Fz1=$fz11 \t $iteration \t  $tot"

exit;
```

ICCP^{3}M procedure for two oposite charges inside a dielectric cylinder is shown in Code Box 11. The cascading procedure, which repeats P^{3}M optimization and ICCP^{3}M until the induced charge values reach a stable value. This cascading procedure has explained and explored in Section 4.6.4.

A.4 Induced Charge Computation

The portion of the ICC code that computes the induced charge values on the cylindrical dielectric interface is shown here that utilize GSL library [213] functions in Code Box 12.

Code Box 12 *ICC Matrix Equation: Solving for the Induced Charges*

```
   /*

      The Core function of the ICC code:
      Polarization Charge Density h determined by LU decomposition.

    */

   void LU_get_h() {
   int s,i,j,k;
   double *workA;
   double totalgrids;
   totalgrids=icylinder.mgrids*icylinder.mgrids;
       iccmatrix.h=malloc(totalgrids*sizeof(double));
      /*
         Temproray array for Matrix A:
         Induced charge-induced charge interactions
       */
       workA=malloc(totalgrids*totalgrids*sizeof(double));
            k=0;
           for(i=0;i<totalgrids;i++) { /* form work A*/
            for(j=0;j<totalgrids;j++) {
              workA[k]=iccmatrix.A[i][j];
              k++;
            }
           }
    /* GSL structures and function to solve the matrix equation Ah=c */
    gsl_matrix_view m = gsl_matrix_view_array (workA, totalgrids, totalgrids);
    gsl_vector_view b = gsl_vector_view_array (iccmatrix.C, totalgrids);
    gsl_vector *h = gsl_vector_alloc (totalgrids);

     gsl_permutation * p = gsl_permutation_alloc (totalgrids);
     gsl_linalg_LU_decomp (&m.matrix, p, &s);
     gsl_linalg_LU_solve (&m.matrix, p, &b.vector, h);

      /* Assigne the solution: Induced charges on the cylinder */
      for(i=0;i<totalgrids;i++) {
       iccmatrix.h[i]=gsl_vector_get(h,i);
        /* printf("h[%d] = %f\n",i,iccmatrix.h[i]);*/
      }
     gsl_vector_fprintf (stdout, h, "%g");
    /* free work arrays */
     gsl_permutation_free(p);
     free(workA);
}
```

The main structure of ICC implementation for dielectric cylinder is summarized in Code Box 13. Results obtained from this code has been compared with ICCP^{3}M result in Section 4.7.

Code Box 13 *C Structures of ICC for dielectric cylinder*

```
typedef struct { /* cylinder/grid data : (phi,rho,z) */
 double radius,e1,e2;
 double **phi,*phi1; /* Coordinates in cylindrical
                      coordinates for grid ids as 2d array */
 double **z,*z1;     /* phi1 and z1 are 1 d arrays */
 int mgrids; /* number of grids in 1 line , actual mgrids x mgrids */
 int select; /* select whole scheme */
 int density,energy,potential,force; /* select output */
 double density_point,start,end,resolution; /* points on the z axis (0,0,z) */
 int c_sum_tr,L_sum_tr;
 double c_int_tr,L_int_tr; /* integrals appear in the ICC formulation */
 double length;
 double area;
 int green;
 double h_phi;  /* printing out density along a constant phi */
 int h_print;
} cylinder;


typedef struct { /* Matrix equation thing */
 double **A,*C,*h;
} matrix_data;
```

B Bibliography

[1] S. Chu. Biology and polymer physics at the single-molecule level. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 361(1805):689–698, 2003.

[2] F. Ritort. Single-molecule experiments in biological physics: methods and applications. *J. Phys.: Cond. Matt.*, 18(32):R531–R583, 2006.

[3] P.V. Cornish and T. Ha. A survey of single-molecule techniques in chemical biology. *ACS Chemical Biology*, 2:53, 2006.

[4] B. Hille. *Ionic channels of excitable membranes*. Sinauer Associates Sunderland, Mass, 1992.

[5] Amit Meller. Dynamics of polynucleotide transport through nanometer-scale pores. *J. Phys.: Cond. Matt.*, 15:R581–R607, 2003.

[6] C. Dekker. Solid-state nanopores. *Nature Nanotech*, 2:209–215, 2007.

[7] Reto B. Schoch, Jongyoon Han, and Philippe Renaud. Transport phenomena in nanofluidics. *Reviews of Modern Physics*, 80(3):839, 2008.

[8] A. Engel and D.J. Müller. Observing single biomolecules at work with the atomic force microscope. *Nature Struct. Bio.*, 7:715–718, 2000.

[9] R. Merkel. Force spectroscopy on single passive biomolecules and single biomolecular bonds. *Physics Reports*, 346(5):343–385, 2001.

[10] TR Strick, MN Dessinges, G. Charvin, NH Dekker, JF Allemand, D. Bensimon, and V. Croquette. Stretching of macromolecules and proteins. *Reports on Progress in Physics*, 66(1):1–45, 2003.

[11] F. Ritort. Work fluctuations, transient violations of the second law and free-energy recovery methods: Perspectives in theory and experiments. *Seminaire Poincare*, 2:193, 2004.

[12] C. Bustamante, J. Liphardt, and F. Ritort. The nonequilibrium thermodynamics of small systems. *Physics Today*, 58(7):43–48, 2005.

[13] P. Gaspard. Out-of-equilibrium nanosystems. *Progress of Theoretical Physics Supplement*, 2006.

[14] P. Gaspard. Hamiltonian dynamics, nanosystems, and nonequilibrium statistical mechanics. *Physica A: Statistical Mechanics and its Applications*, 369(1):201–246, 2006.

[15] S. Sasa. Effective description of small non-equilibrium systems. *Progress of Theoretical Physics Supplement*, 2006.

[16] C. Jarzynski. Work fluctuation theorems and single-molecule biophysics (physics of non-equilibrium systems: Self-organized structures and dynamics far from equilibrium). *Progress of theoretical physics. Supplement*, 165:1–17, 2006.

[17] D. Keller, D. Swigon, and C. Bustamante. Relating single-molecule measurements to thermodynamics. *Biophys. J.*, 84(2):733–738, 2003.

[18] J.F. Marko and E.D. Siggia. Stretching dna. *Macromolecules*, 28(26):8759–8770, 1995.

[19] S. Sinha and J. Samuel. Inequivalence of statistical ensembles in single molecule measurements. *Phys. Rev. E*, 71(2):21104, 2005.

[20] E. Van der Straeten and J. Naudts. A one-dimensional model for theoretical analysis of single molecule experiments. *J. Phys. A: Math. Gen*, 39:5715–5726, 2006.

[21] C. Guardiani and F. Bagnoli. A toy model of polymer stretching. *J. Chem. Phys.*, 125:084908, 2006.

[22] F. Hanke and H.J. Kreuzer. Nonequilibrium theory of polymer stretching based on the master equation. *Phys. Rev. E*, 72(3):31805, 2005.

[23] F. Hanke, H.J. Kreuzer, et al. Nonequilibrium dynamics of single polymer molecules: Relaxation close to and far from equilibrium. *International Journal of Quantum Chemistry*, 106(14):2953–2959, 2006.

[24] HJ Kreuzer, SH Payne, and L. Livadaru. Stretching a macromolecule in an atomic force microscope: Statistical mechanical analysis. *Biophys. J.*, 80(6):2505–2514, 2001.

[25] HJ Kreuzer and SH Payne. Stretching a macromolecule in an atomic force microscope: Statistical mechanical analysis. *Phys. Rev. E*, 63(2):21906, 2001.

[26] M. Zemanová and T. Bleha. Isometric and isotensional force-length profiles in polymethylene chains. *Macromolecular Theory and Simulations*, 14(9):596, 2005.

[27] J.H. Weiner. *Statistical Mechanics of Elasticity*. Dover Publications, 2002.

[28] R.M. Neumann. Nonequivalence of the stress and strain ensembles in describing polymer-chain elasticity. *Phys. Rev. A*, 31(5):3516–3517, 1985.

[29] R.M. Neumann. Implications of using the entropy spring model for an ideal polymer chain. *Phys. Rev. A*, 34(4):3486–3488, 1986.

[30] R.M. Neumann. On the precise meaning of extension in the interpretation of polymer-chain stretching experiments. *Biophys. J.*, 85(5):3418–3420, 2003.

[31] Stijn van Dorp Diego Krapf Ralph M. M. Smeets Serge G. Lemay Nynke H. Dekker Ulrich F. Keyser, Bernard N. Koeleman and Cees Dekker. Direct force measurements on dna in a solid-state nanopore. *Nature Physics*, 2:473 – 477, 2006.

[32] UF Keyser, J. van der Does, C. Dekker, and NH Dekker. Optical tweezers for force measurements on dna in nanopores. *Review of Scientific Instruments*, 77:105105, 2006.

[33] R.D. Astumian. Equilibrium theory for a particle pulled by a moving optical trap. *J. Chem. Phys.*, 126:111102, 2007.

[34] J. Zhang and BI Shklovskii. Effective charge and free energy of dna inside an ion channel. *Phys. Rev. E*, 75(2):21906, 2007.

[35] K. Kawai and K. Okumura. Single molecular statistics of an optically tweezed polymer: A theoretical consideration. *Chemical Physics Letters*, 439(4-6):369–373, 2007.

[36] M. Karttunen, I. Vattulainen, and A. Lukkarinen. *Novel Methods in Soft Matter Simulations*, volume 640 of *Lecture Notes in Physics*. Berlin Springer Verlag, 2004.

[37] S.O. Nielsen, C.F. Lopez, G. Srinivas, and M.L. Klein. Coarse grain models and the computer simulation of soft materials. *Journal of Physics Condensed Matter*, 16(15):481–512, 2004.

[38] Pep Espnol. Statistical mechanics of coarse-graining. *Editors Karttunen, M. and Lukkarinen, A. and Vattulainen, I., Novel Methods in Soft Matter Simulations, Lecture Notes in Physics, Berlin Springer Verlag*, 640:69–115, 2004.

[39] M.J. Stevens. Simple simulations of dna condensation. *Biophys. J.*, 80(1):130–139, 2001.

[40] Markus Deserno. Counterion condensation for rigid linear polyelectrolytes. *Ph.D. Thesis Johannes Gutenberg-Universitaet*, 2000.

[41] Adrian Parsegian. Energy of an ion crossing a low dielectric membrane: Solutions to four relevant electrostatic problems. *Nature*, 221:844, 1969.

[42] Boaz Nadler, Uwe Hollebach, and R.S. Eisenberg. Dielectric boundary force and its crucial role in gramicidin. *Phy. Rev. E*, 68:021905, 2003.

[43] Andrij E. Yaroshchuk. Dielectric exclusion of ions from membranes. *Advances in Colloid and Interface Science*, 85:193–230, 2000.

[44] Dezso Boda, Monika Valisko, Bob Eisenberg, Wolfgang Nonner, Douglas Henderson, and Dirk Gillespie. The effect of protein dielectric coefficient on the ionic selectivity of a calcium channel. *J. Chem. Phys.*, 125(3):034901, 2006.

[45] Barry Honig and Anthony Nicholls. Classical electrostatics in biology and chemistry. *Science*, 268(5214):1144–1149, 1995.

[46] DG Levitt. Electrostatic calculations for an ion channel. i. energy and potential profiles and interactions between ions. *Biophys. J.*, 22(2):209–219, 1978.

[47] P. B. Shaw. Theory of the poisson green's function for discontinuous dielectric media with an application to protein biophysics. *Phys. Rev. A*, 32(4):2476–2487, 1985.

[48] Rosalind Allen, Jean Pierre Hansen, and Simone Melchionna. Electrostatic potential inside ionic solution confined by dielectrics: a variational approach. *Phys. Chem. Chem. Phys.*, 3:4177–4186, 2001.

[49] BZ Lu, YC Zhou, MJ Holst, and JA McCammon. Recent progress in numericalmethods for the poisson-boltzmann equation in biophysical applications. *Communications In Computational Physics*, 3(5):973–1009, 2008.

[50] RD Coalson and TL Beck. Numerical methods for solving poisson and poisson-boltzmann type equations. *Encyclopedia of Computational Chemistry. P. v. R. Schleyer, NL Allinger, T. Clark, J. Gasteiter, PA Kollman, HF Schaefer, III, and HF Schreiner, editors*, 3:2080–2100, 1998.

[51] Rangahathan Bharadwaj, Andreas Windemuth, S.Sridharan, Barry Honig, and Anthony Nicholls. A fast multipole boundary element method for molecular electrostatics: An optimal approach for large systems. *J. Comp. Chem.*, 16(7):898–913, 1995.

[52] D.M. Young. A historical overview of iterative methods. *Computer Physics Communications*, 53(1-3):1–17, 1989.

[53] Jesus A. Izaguirre, Scott S.Hampton, and Thierry Matthey. Parallel multigrid summation for the n-body problem. *J.Parallel Distrib.Comput.*, 65:949–962, 2005.

[54] I.C. Waring, N.Rooney, A.Stewart, and V.F.Fusco. The parallel computation of solutions to electrostatic problems using multigrid techniques. *International Journal of Numerical Modelling:Electrinic Networks,Devices and Fields*, 7:69–74, 1994.

[55] Taro Sometani and Kenroku Hasebe. Method for solving electrostatic problems having a simple dielectric boundary. *Am. Jour. Phys.*, 45:918, 1977.

[56] T.P.Doerr and Yi-Kuo Yu. Electrostatics in the presence of dielectrics: The benefits of treating the induced surface charge density directly. *Am.J.Phys.*, 72:190, 2004.

[57] Rosalind Allen and Jean-Pierre Hansen. Density functional approach to the effective interaction between charges within dielectric cavities. *J. Phys.: Condens.Matter*, 14:11981–11997, 2002.

[58] Dezso Boda, Dirk Gillespie, Wolfgang Nonner, Douglas Henderson, and Bob Eisenberg. Computing induced charges in inhomogeneous dielectric media: Application in a monte carlo simulation of complex ionic system. *Phy. Rev. E*, 69:046702, 2004.

[59] RJ Zauhar and RS Morgan. A new method for computing the macromolecular electric potential. *J Mol Biol*, 186(4):815–20, 1985.

[60] Markus Deserno and Christian Holm. How to mesh up ewald sums. i a theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.*, 109(18):7678, 1998.

[61] H.J. Limbach, A. Arnold, B.A. Mann, and C. Holm. Espresso–an extensible simulation package for research on soft matter systems. *Computer Physics Communications*, 174:704–727, 2006.

[62] M. Rubinstein and R.H. Colby. *Polymer physics*. Oxford University Press, 2008.

[63] T. Keii. *Kinetics of Ziegler-Natta Polymerization*. Kluwer Academic Publishers, 1972.

[64] B. Sakmann, J. Patlak, and E. Neher. Single-channel currents recorded from membrane of denervated frog muscle fibers. *Nature*, 260:799–802, 1976.

[65] G. Binning, CF Quate, C. Gerber, et al. Atomic force microscope. *Phys. Rev. Lett*, 56(9):930–933, 1986.

[66] P.G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, 1979.

[67] H. Schiessel. The physics of chromatin. *Journal Of Physics Condensed Matter*, 15(19):699–774, 2003.

[68] AG Leslie, S. Arnott, R. Chandrasekaran, and RL Ratliff. Polymorphism of dna double helices. *J Mol Biol*, 143(1):49–72, 1980.

[69] A. Kornberg and TA Baker. *DNA replication. WH*. Freeman and Company New York, N. Y, 1992.

[70] A. Ashkin, J.M. Dziedzic, J.E. Bjorkholm, and S. Chu. Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*, 11(5):288–290, 1986.

[71] S. Chu. Laser manipulation of atoms and particles. *Science*, 253(5022):861–866, 1991.

[72] M.D. Wang, H. Yin, R. Landick, J. Gelles, and S.M. Block. Stretching dna with optical tweezers. *Biophys. J.*, 72(3):1335–1346, 1997.

[73] J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco, and C. Bustamante. Reversible unfolding of single rna molecules by mechanical force. *Science*, 292(5517):733–737, 2001.

[74] J.D. Wen, M. Manosas, P.T.X. Li, S.B. Smith, C. Bustamante, F. Ritort, and I. Tinoco Jr. Force unfolding kinetics of rna using optical tweezers. i. effects of experimental variables on measured results. *Biophys. J.*, 92(9):2996, 2007.

[75] Jeffrey R. Moffitt, Yann R. Chemla, Steven B. Smith, and Carlos Bustamante. Recent advances in optical tweezers. *Annual Review of Biochemistry*, 77(1):205–228, 2008.

[76] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, and H.E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by afm. *Science*, 276(5315):1109, 1997.

[77] F. Oesterhelt, M. Rief, and HE Gaub. Single molecule force spectroscopy by afm indicates helical structure of poly (ethylene-glycol) in water. *New J. Phys*, 1(6.1), 1999.

[78] T.E. Fisher, P.E. Marszalek, and J.M. Fernandez. Stretching single molecules into novel conformations using the atomic force microscope. *Nature Struct. Bio.*, 7:719–724, 2000.

[79] MI Giannotti and GJ Vancso. Interrogation of single synthetic polymer chains and polysaccharides by afm-based force spectroscopy. *Chemphyschem*, 8(16):2290–307, 2007.

[80] SB Smith, L. Finzi, and C. Bustamante. Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads. *Science*, 258(5085):1122–1126, 1992.

[81] TR Strick, J.F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The elasticity of a single supercoiled dna molecule. *Science*, 271(5257):1835, 1996.

[82] K. Besteman, S. Hage, NH Dekker, and SG Lemay. Role of tension and twist in single-molecule dna condensation. *Phys. Rev. Letters*, 98(5):58103, 2007.

[83] T.T. Perkins, D.E. Smith, and S. Chu. Single Polymer Dynamics in an Elongational Flow. *Science*, 276(5321):2016, 1997.

[84] R.M. Neumann. Polymer stretching in an elongational flow. *J. Chem. Phys.*, 110:7513, 1999.

[85] S. Weiss. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature Struct. Bio.*, 7:724–729, 2000.

[86] K. Huang. *Statistical Mechanics*. Wiley, New York, 1987.

[87] M.J. Lang and S.M. Block. Resource letter: Lbot-1: Laser-based optical tweezers. *Am. J. Phys.*, 71:201, 2003.

[88] M. Daune. *Molecular Biophysics: Structures in Motion*. Oxford University Press, Oxford, 1999.

[89] C.R. Cantor and P.R. Schimmel. *The behavior of biological macromolecules*. WH Freeman, 1980.

[90] Bob Eisenberg. Ionic channels as natural nanodevices. *J. Comp. Electronics*, 1:331–333, 2002.

[91] Benoit Roux. Theoretical and computational models of ion channels. *Curr. Opin. Struct. Biol.*, 12:182–189, 2002.

[92] Shin-Ho Chung and D. Peter Tielman. Computational and theoretical approached to unraveling the permeation dynamics in biological nanotubes. *Handbook of Theoretical and Computational Nanotechnology*, X:1–54, 2005.

[93] Boaz Nadler, Zeev Schuss, Amit Singer, and R S Eisenberg. Ionic diffusion through confined geometries: from langevin equations to partial differential equations. *J. Phys.: Cond. Matt.*, 16:S2153–S2165, 2004.

[94] John J. Kasianowicz, Eric Brandin, Daniel Branton, and David W. Deamer. Characterization of individual polynucleotide molecules using a membrane channel. 93:13770–13773, 1996.

[95] B. Roux and M. Karplus. Ion transport in a gramicidin-like channel: dynamics and mobility. *The Journal of Physical Chemistry*, 95(12):4856–4868, 1991.

[96] B. Roux and M. Karplus. Ion transport in a model gramicidin channel. structure and thermodynamics. *Biophysical Journal*, 59(5):961–981, 1991.

[97] Simon Berneche and Benoit Roux. Energetics of ion conduction through the k+ channel. *Nature*, 414:73–77, 2001.

[98] Shin-Ho Chung and Serdar Kuyucak. Recent advances in ion channel research. *Biochim. Biophys.*, 1565:267–286, 2002.

[99] Wonpin Im and Benoit Roux. Ion permeation and selectivity of ompf porin: A theoretical study bases in molecular dynamics,brownian dynamics and continuum electrodiffusion theory. *J.Mol.Biol.*, 312:851–869, 2002.

[100] Ben Corry, Serdar Kuyucak, and Shin-Ho Chung. Dielectric self-energy in poisson-boltzmann and poisson-nernst-planck models of ion channels. *Biophy. J.*, 84:3594–3606, 2003.

[101] John J. Kasianowicz, Sarah E. Henrickson, Howard H. Weetall, and Baldwin Robertson. Simultaneous multianalyte detection with a nanometer-scale pore. *Anal.Chem.*, 73:2268–2272, 2001.

[102] Lucas Nivon Amit Meller and Daniel Branton. Voltage-driven dna translocations through a nanopore. *Phys. Rev. Lett.*, 86(15):3435, 2001.

[103] David K. Lubensky and David R.Nelson. Driven polymer translocation through a narrow pore. *Biophy. J.*, 77:1824–1838, 1999.

[104] M.Muthukumar. Polymer translocation through a hole. *J. Chem. Phys.*, 111(22):10371, 1999.

[105] Pierre-Gilles de Gennes. Passive entry of a dna molecule into a small pore. 76:7262–7264, 1999.

[106] K.L.Sebastian and Alok K.R. Paul. Kramers problem for a polymer in a double well. *Phy. Rev. E*, 62(1):927, 2000.

[107] Jeffrey Chuang, Yacov Kantor, and Mehran Kardar. Anomalous dynamics of translocation. *Phy. Rev. E*, 65:011802, 2001.

[108] Z.Konkoli E.A.Di Marzio T.Ambjomsson, S.P.Apell and J.J.Kasianowicz. Charged polymer membrane translocation. *J. Chem. Phys.*, 117(8):4063, 2002.

[109] Yacov Kantor and Mehran Kardar. Anomalous dynamics of forced translocation. *Phy. Rev. E*, 69:021806, 2004.

[110] Rhonald C. Lua and Alexander Y. Grosberg. First passage time and asymmetry of dna translocation. *Phy. Rev. E*, 72:061918, 2005.

[111] H. Vocks, D. Panja, G.T. Barkema, and R.C. Ball. Pore-blockade times for field-driven polymer translocation. *Journal of Physics: Condensed Matter*, 20:095224, 2008.

[112] Shyh-Shi Chern, Alfredo E. Cardenas, and Rob D. Coalson. Three-dimensional dynamic monte carlo simulation of driven polymer through a hole in a wall. *J. Chem. Phys.*, 115(16):7772, 2001.

[113] Yves Lansac, Prabal K. Maiti, and Matthew A. Glaser. Coarse-grained simulation of polymer translocation through an artificial nanopore. *Polymer*, 45:3099–3110, 2004.

[114] Yitzhak Rabin and Motoniko Tanaka. Dna in nanopores: Counterion condensation and coion depletion. *Phys. Rev. Lett.*, 94:148103, 2005.

[115] Kaifu Luo, T. Ala-Nissila, and See-Chen Ying. Polymer translocation through a nanopore: A two-dimensional monte carlo study. *J. Chem. Phys.*, 124:034714, 2006.

[116] Kaifu Luo, T. Ala-Nissila, and See-Chen Ying. Polymer translocation through a nanopore under an applied external field. *J. Chem. Phys.*, 124:114704, 2006.

[117] Matteo Pasquali Anatoly B. Kolomeisky Silvina Matysiak, Alberto Montesi and Cecilia Clementi. Dynamics of polymer translocation through nanopores: Theory meets experiment. *Phys. Rev. Lett.*, 96:118103, 2006.

[118] M. Muthukumar and C.Y.Kong. Simulation of polymer translocation through protein channels. *PNAS USA*, 103:5273–5278, 2006.

[119] A.J.Storm, J.H. Chen, H.W.Zandbergen, and C.Dekker. Translocation of double-strand dna through a silicon oxide nanopore. *Phy. Rev. E*, 71:051903, 2005.

[120] Ralph M. M. Smeets, Ulrich F. Keyser, Diego Krapf, Meng-Yue Wu, Nynke H. Dekker, and Cees Dekker. Salt dependence of ion transport and dna translocation through solid-state nanopores. *Nano Letters*, 6:89–95, 2006.

[121] Derek Stein, Maarten Kruithof, and Cees Dekker. Surface-charge-governed ion transport in nanofluidic channels. *Phys. Rev. Lett.*, 93:035901, 2004.

[122] I.D. Kosinska A. Fuliski and Z.Siwy. On the validity of continuos modelling of ion transport through nanochannels. *Europhys.Lett.*, 67:683–689, 2004.

[123] E.A. Lipman, B. Schuler, O. Bakajin, and W.A. Eaton. Single-molecule measurement of protein folding kinetics. *Science*, 301(5637):1233–1235, 2003.

[124] P. J. Flory. *Statistical Mechanics of Chain Molecules*. Hanser Publishing, Munich, 1989.

[125] E. Guth and H. Mark. Zur innermolekularen, statistik, insbesondere bei kettenmolekiilen i. *Monatshefte für Chemie/Chemical Monthly*, 65(1):93–121, 1934.

[126] P. Pincus. Excluded volume effects and stretched polymer chains. *Macromolecules*, 9(3):386–388, 1976.

[127] D. Perchak and J.H. Weiner. Classical and quantum calculations for short stretched chain models. *Macromolecules*, 15(2):545–549, 1982.

[128] J.H. Weiner. Use of s k log p for stretched polymers. *Macromolecules*, 15(2):542–544, 1982.

[129] JH Weiner. Entropic versus kinetic viewpoints in rubber elasticity. *Am. J. Phys.*, 55:746, 1987.

[130] RG Winkler and P. Reineker. Finite size distribution and partition functions of Gaussian chains: maximum entropy approach. *Macromolecules*, 25(25):6891–6896, 1992.

[131] D.E. Makarov, Z. Wang, J.B. Thompson, and H.G. Hansma. On the interpretation of force extension curves of single protein molecules. *J. Chem. Phys.*, 116:7760, 2002.

[132] R.G. Winkler. Deformation of semiflexible chains. *J. Chem. Phys.*, 118:2919, 2003.

[133] A.M. Skvortsov, L.I. Klushin, and FA Leermakers. On the escape transition of a tethered gaussian chain; exact results in two conjugate ensembles. *Macromolecular Symposia*, 237:73, 2006.

[134] I. Webman, J.L. Lebowitz, and MH Kalos. Elastic properties of a polymer chain. *Phys. Rev. A*, 23(1):316–320, 1981.

[135] JH Weiner and D. Perchak. Frenkel's governor model for stretched polymers. *Macromolecules*, 14(5):1590–1591, 1981.

[136] JH Weiner and DH Berman. Bond forces in long-chain molecules. *J. Chem. Phys.*, 82:548, 1985.

[137] DH Berman and JH Weiner. Compressive axial forces in long-chain molecules due to excluded-volume effects. *J. Chem. Phys.*, 83:1311, 1985.

[138] RA Guyer and Jay Johnson. Equations of state of a single polymer chain. *Phys. Rev. A*, 32(6):3661–3664, 1985.

[139] P. Cifra and T. Bleha. Force–extension relations in macromolecules of variable excluded volume and flexibility: energy and entropy changes on stretching. *J. Chem. Soc., Faraday Trans*, 91:2465–2471, 1995.

[140] JT Titantah, C. Pierleoni, and J.P. Ryckaert. Different statistical mechanical ensembles for a stretched polymer. *Phys. Rev. E*, 60(6):7010–7021, 1999.

[141] R. P. Linna and K. Kaski. Analysis of dna elasticity. *Phys. Rev. Letters*, 100(16):168104, 2008.

[142] Daniel F. Styer. What good is the thermodynamic limit? *Am. J. Phys.*, 72(1):25–29, 2004.

[143] J.P. Hansen and I.R. McDonald. *Theory of Simple Liquids*. Academic Press, 2002.

[144] H. Touchette, R.S. Ellis, and B. Turkington. An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles. *Physica A: Statistical Mechanics and its Applications*, 340(1-3):138–146, 2004.

[145] M. Doi and S.F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, 1986.

[146] R. Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255–284, 1966.

[147] Malcolm E. Davis and J. Andrew Mccammon. Electrostatics in biomolecular structure and dynamics. *Chem.Rev.*, 90:509–521, 1990.

[148] R.S.Eisenberg. Computing the field in proteins and channels. *J. Membrane Biol.*, 150:1–25, 1996.

[149] M.J. Holtz. The poisson-bolztmann equation: Analysis and multilevel solution. *Ph.D. Thesis University of Illinois Urbana-Champaign*, 1994.

[150] Mike P. Allen and Dominik J. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications. Clarendon Press, Oxford, 1 edition, 1987.

[151] R. W. Hockney and J. W. Eastwood. Computer simulations using particles. 1981.

[152] B. Roux, B. Prod'hom, and M. Karplus. Ion transport in the gramicidin channel: molecular dynamics study of single and double occupancy. *Biophys. J.*, 68(3):876–892, 1995.

[153] J. Warwicker and HC Watson. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J Mol Biol*, 157(4):671–9, 1982.

[154] ME Davis and JA McCammon. Solving the finite difference linearized Poisson-Boltzmann equation: A comparison of relaxation and conjugate gradient methods. *Journal of Computational Chemistry*, 10(3):386–391, 1989.

[155] A. Nicholls and B. Honig. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the poisson-boltzmann equation. *J. Comput. Chem*, 12(4):435–445, 1991.

[156] Maria G. Kornikova, Rob D. Coalson, Peter Graf, and Abraham Nitzan. A lattice relaxation algorithm for three-dimensional posisson-nernst-planck theory with application to ion transport through the gramicidin a channel. *Biophy. J.*, 76:642–656, 1999.

[157] A.A. Rashin. Hydration phenomena, classical electrostatics, and the boundary element method. *The Journal of Physical Chemistry*, 94(5):1725–1733, 1990.

[158] S. Grandison, R. Penfold, and J.M. Vanden-Broeck. Monte carlo simulation of an inhomogeneous dielectric continuum model for b-dna. *Phys. Chem. Chem. Phys.*, 7:3486 – 3495, 2005.

[159] T. P. Doerr and Yi-Kuo Yu. Electrostatics of charged dielectric spheres with application to biological systems. *Phys. Rev. E (Statistical, Nonlinear, and Soft Matter Physics)*, 73(6):061902, 2006.

[160] T.J. You and S.C. Harvey. Finite element approach to the electrostatics of macromolecules with arbitrary geometries. *Journal of Computational Chemistry*, 14(4):484–501, 1993.

[161] Massiomo Marchi, Daniel Borgis, Nicolas Levy, and Pietro Ballone. A dielectric continuum molecular dynamics method. *J. Chem. Phys.*, 114(10):4377, 2001.

[162] Phil Attard. Variational formulation for the electrostatic potential in dielectric continua. *J. Chem. Phys.*, 119:1365, 2003.

[163] Douglas Henderson David Busath and Stefan Sokolowski. Density functional theory for an electrolyte in a cylinder: the selectivity of a calcium channel. *J.Phys.: Condens.Matter*, 16:S2193–S2201, 2004.

[164] Dezso Boda, Tibor Varga, Douglas Henderson, David D. Busath, Wolfgang Nonner, Dirk Gillespie, and Bob Eisenberg. Monte carlo simulation study of a system with a dielectric boundary: Application to calcium channel selectivity. *Molecular Simulation*, 30:89–96, 2004.

[165] Nicolas Levy, Daniel Borgis, and Massimo Marchi. A dielectric continuum model of solvation of complex solutes. *Comp.Phys.Comm.*, 169:69–74, 2005.

[166] Rosa Ramirez, Ralph Gebauer, and Michel Mareschal. Density functional theory of solvation in a polar solvent: Extracting the functional from homegenous solvent simulations. *Phy. Rev. E*, 66:031206, 2002.

[167] Rose Ramirez and Daniel Borgis. Density functional theory of solvation and its relation to implicit solvent models. *J.Phys.Chem.B*, 109:6754–6763, 2005.

[168] Nicolas Levy, Daniel Borgis, and Massimo Marchi. Embedding a molecular solute in a macroscopic solvent: A dielectric continuum molecular dynamics method. *Simu*, Issue 5,Chapter IV.

[169] P.G.Martinsson and V.Rokhlin. A fast direct solver for boundary integral equations in two dimensions. *J.Comp.Phys.*, 205:1–23, 2005.

[170] Per-Gunnar Martinsson. Fast evaluation of electro-static interactions in multi-phase dielectric media. *J.Comp.Phys.*, 211:289–299, 2006.

[171] A. Duncan, R. D. Sedgewick, and R. D. Coalson. Local simulation algorithms for coulomb gases with dynamical dielectric effects. *Phys. Rev. E (Statistical, Nonlinear, and Soft Matter Physics)*, 73(1):016705, 2006.

[172] A. Duncan and R. D. Sedgewick. Accelerated multiboson algorithm for coulomb gases with dynamical dielectric effects. *Phys. Rev. E (Statistical, Nonlinear, and Soft Matter Physics)*, 73(6):066711, 2006.

[173] T. Biben, J. P. Hansen, and Y. Rosenfeld. Generic density functional for electric double layers in a molecular solvent. *Phys. Rev. E*, 57(4):R3727–R3730, Apr 1998.

[174] Daniel M. Chipman. New formulation and implementation for volume polarization in dielectric continuum theory. *J. Chem. Phys.*, 124(22):224111, 2006.

[175] Sofia Teber. Translocation energy of ions in nano-channels of cell membranes. *J. Stat. Mech.*, page P07001, 2005.

[176] J. D. Jackson. *Classical Electrodynamics*. Wiley, New York, 3rd edition, 1999.

[177] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical recipes in c: The art of scientific computing. 1992.

[178] W. Kahan. *Gauss-Seidel methods of solving large systems of linear equations*. PhD thesis, University of Toronto, 1958.

[179] George B Arfken and Hans J. Weber. Mathematical methods for physicists. Academic Press, 1995.

[180] Charles S.Peskin Sanford M.Simon and George F. Oster. What drives the translocation of proteins. *Proc. Natl. Acad. Sci. USA*, 89:3770–3774, 1992.

[181] W. Sung and P.J. Park. Polymer translocation through a pore in a membrane. *Phys. Rev. Lett.*, 77(4):783, 1996.

[182] Serdar Kuyucak and Turgut Bastug. Physics of ion channels. *J. Biol. Phys.*, 29:429–446, 2003.

[183] Gennady V. Miloshevsky and Peter C. Jordan. Permeation in ion channels: the interplay of structure and theory. *TRENDS in Neuroscience*, 27:308, 2004.

[184] Simon Berneche Benoit Roux, Toby Allen and Wonpil Im. Theoretical and computational models of biological ion channels. *Quarterly Reviews og Biophysics*, 37:15–103, 2004.

[185] Roland Roth and Dirk Gillespie. Physics of size selectivity. *Phys. Rev. Lett.*, 95:247801, 2005.

[186] F.G.Ball and S.P.Samsom. Ion-channel gating mechanisms: Model identification and parameter estimation from single channel recordings. *Proc.Roy.Soc.Lon. Ser. B, Biological Sci.*, 236:385–416, 1989.

[187] J. Li, X. Gong, H. Lu, D. Li, H. Fang, and R. Zhou. Electrostatic gating of a nanometer water channel. *Proceedings of the National Academy of Sciences*, 104(10):3687, 2007.

[188] Serdar Kuyucak, Olaf Sparre Andersen, and Shin-Ho Chung. Models of permeation in ion channels. *Rep. Prog. Phys.*, 64:1427–1472, 2001.

[189] A. Yu. Grosberg, S. Nechaev, M. Tamm, and O. Vasilyev. How long does it take to pull an ideal polymer into a small hole? *Phys. Rev. Letters*, 96(22):228105, 2006.

[190] A.B. Kolomeisky. Channel-facilitated molecular transport across membranes: Attraction, repulsion, and asymmetry. *Physical review letters*, 98(4), 2007.

[191] D. Panja, G.T. Barkema, and R.C. Ball. Fast track communication: Anomalous dynamics of unbiased polymer translocation through a narrow pore. *Journal of Physics: Condensed Matter*, 19(43):432202, 2007.

[192] C.T.A. Wong and M. Muthukumar. Polymer translocation through a cylindrical channel. *J. Chem. Phys.*, 128:154903, 2008.

[193] P.T. Underhill and P.S. Doyle. On the coarse-graining of polymers into bead-spring chains. *J. Non-Newtonian Fluid Mech*, 122(1):3–31, 2004.

[194] M. Peyrard and AR Bishop. Statistical mechanics of a nonlinear model for dna denaturation. *Phys. Rev. Letters*, 62(23):2755–2758, 1989.

[195] Michel Peyrard. Nonlinear dynamics and statistical physics of dna. *Nonlinearity*, 17(2):R1–R40, 2004.

[196] C. Bouchiat, M. D. Wang, J.-F. Allemand, T. Strick, S. M. Block, and V. Croquette. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.*, 76(1):409–413, 1999.

[197] J. Wang and H. Gao. A generalized bead-rod model for brownian dynamics simulations of wormlike chains under strong confinement. *J. Chem. Phys.*, 123:0e4906, 2005.

[198] G. Wang, B. Zhang, J.R. Wayment, J.M. Harris, and H.S. White. Electrostatic-gated transport in chemically modified glass nanopore electrodes. *J. Am. Chem. Soc*, 128(23):7679–7686, 2006.

[199] T. Dauxois. Dynamics of breather modes in a nonlinear helicoidal model of dna. *Physics letters. A*, 159(8-9):390–395, 1991.

[200] Thierry Dauxois, Michel Peyrard, and A. R. Bishop. Entropy-driven dna denaturation. *Phys. Rev. E*, 47(1):R44–R47, Jan 1993.

[201] F. de los Santos, O. Al Hammal, and M.A. Munoz. Simplified langevin approach to the peyrard-bishop-dauxois model of dna. *Phys. Rev. E*, 77:032901, 2008.

[202] M. Deserno, C. Holm, and S. May. Fraction of condensed counterions around a charged rod: Comparison of poisson-boltzmann theory and computer simulations. *Macromolecules*, 33(1):199–206, 2000.

[203] R. Zandi, D. Reguera, J. Rudnick, and W.M. Gelbart. What drives the translocation of stiff chains? *Proceedings of the National Academy of Sciences*, 100(15):8649–8653, 2003.

[204] L. Livadaru and HJ Kreuzer. Confinement of a polymer chain in a tube. *New Journal of Physics*, 5(1):95, 2003.

[205] Marjolein Dijkstra, Daan Frankel, and Henk N. W. Lekkerkerker. Confinement free energy of semiflexible polymers. *Physica A*, 193:374–393, 2002.

[206] M.O. Khan and D.Y.C. Chan. Monte carlo simulations of stretched charged polymers. *J. Phys. Chem. B*, 107(32):8131–8139, 2003.

[207] M.O. Khan and D.Y.C. Chan. Effect of chain stiffness on polyelectrolyte condensation. *Macromolecules*, 38(7):3017–3025, 2005.

[208] D. Wei, W. Yang, X. Jin, and Q. Liao. Unforced translocation of a polymer chain through a nanopore: The solvent effect. *J. Chem. Phys.*, 126:204901, 2007.

[209] JMG Sarraguca and A. Pais. Polyelectrolytes in solutions with multivalent salt. effects of flexibility and contour length. *Physical Chemistry Chemical Physics*, 8(36):4233–4241, 2006.

[210] Michael D. Altman, Jaydeep P. Bardhan, Jacob K. White, and Bruce Tidor. Accurate solution of multi-region continuum biomolecule electrostatic problems using the linearized poisson-boltzmann equation with curved boundary elements. *Journal of Computational Chemistry*, 30:132–153, 2009.

[211] Jose Nelson Onuchic and Peter G. Wolynes. Theory of protein folding. *Current Opinion in Structural Biology*, 14:70–75, 2004.

[212] Michael Schlierf, Hongbin Li, and Julio M. Fernandez. The unfolding kinetics of ubiquitin captured with single molecule force clamp techniques. *Proceedings of the National Academy of Sciences of the United States of America*, 101(19):7299–7304, May 2004.

[213] GNU Scientific Library. http://www.gnu.org/software/gsl/.

C Acknowledgements

I would like express my graditute for their support and motivation in all these years to my parents Yıldız Süzen and Teoman Süzen†, my brother Ziya Süzen, sister-inlaw Rachel Elisabeth Süzen, my nephews Jason John Süzen and Jamie Robert Süzen, and my niece Romy Süzen, my grandfather UK SBA Insp. M. Kalfağlu † and my aunt Ayla Kalfaoğlu.

I am grateful to Christian Holm for his generous support of this work, letting me work in his group, as well as letting me discover world of coarse-graining and the spirit of modern statistical physics, Marcello Sega for being the source of acceptance of my work within the group and offering his quite precious scientific collaboration and support, Marcia Barbosa for helping me to handle the induced charge computation with density functional approach during its prototype implementation, her kind effort and hospitality during my visit to Brazil and her group's friendly environment, Sandeep Tyagi for sharing his novel idea of utilization of P^3M , and sharing beauty of mathematics, Sofia Kantorovich for her quite careful and critical proof reading of the thesis and fruitful scientific discussions, SimBio Group in general and in particular Kai Grass and Joan Josep Cerdá for their friendly approach and their support, Florian Dommert for his kind help in translating German summary of the thesis and Prakash Kona for his partial English language review in a short notice.

I would like to express my appreciation to FIAS and FIGSS for providing generous financial support and pleasant work environment. DAAD for its generous financial support for my Brazil visit. Frau Walburga Bergmann in FIAS for her friendly face in sorting out administrative issues smoothly.

I would like to thank colleagues and friends for showing friendship and their kind support Ovidiu Florin Jurjut, Joonsuk Huh, Raul Muresan, Raul Vicente, Elsa Henriques and Cornelius Weber. The rest of the people I *must acknowledge* from my previous studies and work before I came to Frankfurt am Main: My previous academic mentors Professor Mustafa Halilsoy and Professor Notker Rösch for their recommendations and encourgements, colleagues and friends in Cyprus Izzet Sakallı, Mehmet Alptürk, Hakan Mavros, M. Erkut Özser, Devrim Seral, Nidai Kordal, A. Tüzünkan, C. Kaanoğlu, Halil Berberoğlu, Yalın Sayı, Tomasz Garstka and Christina Garstka-Hadjitoma for their support in re-discovering my abilities, colleagues and friends in Munich Eberhardt Herdtweck and his Family, Sven Krueger, Monica Fuchs, Walter Alsheimer, Eric le Gentil, Maite Garcia, Mircea Girju, Katcharin Sriwong, Alexei Matveev, Andre Voitersky, Lyudmila Moskaleva, Dmitriy Ganyushin, Alexander Genest, Norge Cruz and Guoming Xiong for their friendly approach, people in Bristol School of Mathematics, Prof. S. Wiggins and his group members, Matt Nunes, Nick Jones, Nuri Ersoy, Geoff Chapman, Rajesh Kumar Handa, Anindia Mohinta, Steve Plasting, Chris Wheeler, Ian Williams, Apala Majumdar, and Illias Kapsis for showing their friendship and support.

Reseach councils DFG, EPSRC and ORS, as well as EMU and Turkish-Cypriot national education authorities for their generous previous financial support before my Frankfurt years.

D Vitae

Mehmet Süzen

Born 20.11.1977, Cypriot

Degrees

September 2000 M.Sc. Physics	Eastern Mediterranean University, Famagusta, Cyprus M.Sc.Thesis Title: "The chaotic behaviour of geodesics in non-homogeneous vacuum pp-wave solutions"
September 1998 B.Sc. Mech. Eng.	ibid.
June 1994 Senior High School	Mathematics Section, 20 Temmuz Senior High School, Nicosia, Cyprus

Background

10/2005 - Present	Postgraduate by Research (with PD Dr. Holm) Frankfurt Institute for Advanced Studies (FIAS) Goethe Universitaet, Frankfurt am Main, Germany
02/2005 - 09/2005	Instructor in Information Systems Engineering Cyprus International University, Nicosia, Cyprus
07/2003 - 01/2005	Internet Systems Engineer/Research Officer Kibris.Net Internet Services, Nicosia, Cyprus
10/2002 - 06/2003	Postgraduate by Research (funded by EPSRC/ORS), Applied Mathematics (Group AMNAG) University of Bristol, Bristol, United Kingdom
12/2000 - 09/2002	Guest Research Asistant (funded by DFG), (with Prof. Rösch) Theoretical Chemistry Technische Universitaet Muenchen, Munich, Germany
08/1998 - 11/2000	Research/Teaching Assistant, Physics (Gravitational Physics Group of Prof. Halilsoy) Eastern Mediterranean University, Famagusta, Cyprus

Professional

Associate Member (AMInstP) of Institute of Physics (IoP)
Member of Association for Computing Machinery (ACM)